Pulp and paper environmental effects monitoring technical guidance document [chap. 11]

[Title supplied by cataloguer]

Table of contents

11. ALTERNATIVE MONITORING METHODS 11-1

11.1 Use of Mescosms as an Alternative Monitoring Method 11-1
11.1.1 Background of Artificial Stream Development 11-1
11.1.2 Applicability within EEM programs 11-5
11.1.2.1 Pulp and Paper EEM 11-5
11.1.2.2 Metal Mining EEM 11-6
11.1.2.2.1 Initial and Periodic Monitoring 11-6
11.1.2.2.2 Focussed Monitoring 11-6
11.1.2.2.3 Investigation of Cause 11-7
11.1.3 Mesocosm Technology 11-7
11.1.3.1 Suitability of Mesocosm Design 11-7
11.1.4 Description and General Operation of Large Transportable Mesocosm System.. 11-7
11.1.5 Considerations for Site Selection 11-14
11.1.6 Biological Monitoring Study Designs 11-15
11.1.6.1 Overview 11-15
11.1.7 Study Treatments 11-16
11.1.8 Replication 11-17
11.1.9 Sample Sizes and the Role of Effect Size 11-18
11.1.9.1 No available pre-existing (historical) data 11-18
11.1.9.2 Available pre-existing (historical) data 11-19
11.1.10 Fish Monitoring –Effects Assessment 11-20
11.1.10.1 Study Design 11-20
11.1.10.1.1 Species Selection 11-20
11.1.10.2 Biological Endpoints 11-21
11.1.10.3 Timing and duration of effluent exposure 11-22
11.1.10.4 Fish Methods 11-23
11.1.10.4.1 General Set-up 11-23
11.1.10.4.2 Maintenance and Monitoring 11-24
11.1.10.4.3 Sampling and Analysis of Fish 11-26
11.1.10.5 Data Assessment and Interpretation 11-26
11.1.10.5.1 General Requirements 11-26
11.1.10.5.2 Statistical Analysis of Parameters 11-27
11.1.10.5.3 Nested ANOVA Analyses 11-29
11.1.10.5.4 ANCOVA Analyses 11-30
11.1.10.6 Data QA/QC and Analysis 11-30
11.1.11 Benthic Invertebrate Community Monitoring Effects Assessment 11-31
11.1.11.1 Study Design 11-31
11.1.11.1.1 Species Selection 11-31
11.1.11.2 Endpoints 11-32
11.1.11.2.1 Biological Endpoints 11-32
11.1.11.3 Timing and duration of effluent exposure 11-33
11.1.11.4 Benthic Invertebrate Methods 11-34
11.1.11.4.1 General Set-up 11-34
11.1.11.4.2 Maintenance and Monitoring 11-35
11.1.11.5 Sampling and Analysis of Algae and Benthic Invertebrates 11-36
11.1.11.6 Data Assessment and Interpretation 11-37
11.1.12 Supporting Measurements 11-38

11.1.12.1 Water Chemistry Endpoints 11-38
11.1.13 Exposure Endpoints 11-39
11.1.14 Other Considerations 11-40
11.1.15 References 11-42

11.2 Use of Caged Bivalves as Alternative Monitoring Method 11-49
11.2.1 Introduction 11-49
11.2.2 Background 11-49
11.2.3 Approach 11-50
11.2.4 Species Selection 11-51
11.2.4.1 Most Commonly Used Taxa 11-52
11.2.4.2 Marine and Estuarine Bivalves 11-54
11.2.4.3 Freshwater Bivalves 11-54
11.2.4.4 Source 11-55
11.2.4.5 Size and Age 11-56
11.2.4.6 Number of Organisms 11-56
11.2.4.7 Handling and Holding Conditions 11-57
11.2.5 Study Design 11-57
11.2.6 Sampling Design 11-58
11.2.6.1 Area and Station Selection 11-59
11.2.7 Replication 11-59
11.2.7.1 Number of Cage Stations 11-59
11.2.7.2 Replication and Pseudoreplication 11-60
11.2.7.3 Applying Replication To Study Design 11-63
11.2.7.4 Timing and Duration 11-65
11.2.7.5 Modifying or Confounding Factors and Supplementary Measures 11-65
11.2.7.6 Cage Designs 11-67
11.2.7.7 Mooring Systems 11-68
11.2.8 Variables to be Measured 11-68
11.2.8.1 Survival 11-72
11.2.8.2 Growth 11-72
11.2.8.3 Condition 11-73
11.2.9 Methods for Test Initiation, Cage Deployment and Retrieval, Test Termination 11-74
11.2.9.1 Test Initiation 11-74
11.2.9.2 Cage Deployment 11-75
11.2.9.3 Cage Retrieval and Test Termination 11-75
11.2.10 Data Analyses 11-76
11.2.11 Tissue Concentrations 11-77
11.2.11.1 Dioxins and Furans 11-77
11.2.11.2 Other Chemicals 11-77
11.2.12 Reporting 11-78
11.2.13 References 11-78

List of Tables

Table 11-1: Summary of artificial stream applications for assessing the effects of pulp and paper and mining effluents on aquatic ecosystems for Canadian Environmental Effects Monitoring 11-3
Table 11-2: Summary of types of information attained from Canadian artificial stream studies in the last decade (after Dubé et al 2001a) 11-4
Table 11-3: Effects on fish. Biological endpoints recommended for determining effects as designated by statistically significant differences between exposure and reference streams. 11-28
Table 11-4: Biological endpoints to be used for supporting analyses. These analyses are for informational purposes and significant differences between exposure and reference streams are not necessarily used to designate an effect. 11-29
Table 11-5: ANOVA Table for Nested Design, with Two treatments, Four replicate Streams per treatment, and *n* Fish or Composite Samples Within Each Replicate Stream 11-30
Table 11-6: ANOVA Table for Repeated Measure (RM) Design with Two Treatments, Four Replicate Streams per treatment, and t Measurements Times 11-39
Table 11-7: Evaluation criteria for EEM monitoring alternatives (after Dubé et al., 2001a).. 11-41
Table 11-8: Suggested Taxa for Use in Caged Bivalve Studies for EEMs 11-53
Table 11-9: An Example of A Field Data Sheet For Recording Survival And Growth Raw Data 11-71
Table 11-10: Growth Measurements 11-72
Table 11-11: Summary of Statistical Procedures to Analyze Effect Endpoints for Caged Bivalve Studies 11-76

List of Figures

Figure 11-1: Picture of mesocosm trailer system used for EEM 11-9
Figure 11-2: Schematic of mesocosm trailer system (not to scale)(after Dubé 2000) 11-10
Figure 11-3: A) mesocosm stream 11-12
Figure 11-4: Fan propeller unit for current velocity 11-13
Figure 11-5: Fish allocation scenarios where A) adults and juveniles are allocated to the same stream or B) adults and juveniles are allocated to separate streams. For actual implementation, streams are randomly allocated to each treatment. 11-25
Figure 11-6: Example of bivalve mooring design 11-70

11. Alternative Monitoring Methods

At some pulp and paper mill sites the fish and benthic invertebrate community monitoring studies may not be appropriate. The reasons for this are site specific, but the most common reasons are the presence of hazardous conditions (e.g. high velocity) or the presence of confounding factors such as other effluent discharges in the exposure area, that will make it difficult or impossible to isolate any effects attributable to the effluent being monitored.

Where mills cannot design the fish or benthic invertebrate community surveys to resolve difficulties associated with confounding influences, mills will provide a scientific rationale and justification and propose cost effective and technically feasible alternative monitoring methods within the study design.
A number of alternative monitoring methods are recommended in this Chapter . Mills may choose other scientifically defensible methods provided that the results can determine if the effluent is having effects on the fish population (growth, reproduction, condition and survival), fish tissue (levels of dioxins and furans in fish tissue) or the benthic invertebrate community (benthic invertebrate density, taxa richness, the Simpson's Diversity Index and the Bray-Curtis Index). Recommended alternative monitoring methods to the fish survey are mesocosm studies, caged bivalves, and expanded sublethal toxicity testing. For benthic invertebrate community surveys, the recommended alternative monitoring method is mesocosm studies.

11.1 Use of Mescosms as an Alternative Monitoring Method

11.1.1 Background of Artificial Stream Development

The environmental effects monitoring (EEM) program was developed to aid in assessing the adequacy of both the *Pulp and Paper Effluent Regulations* and the *Metal Mining Effluent Regulations* under the federal *Fisheries Act*. These programs have been designed to achieve national uniformity in monitoring of effects, while taking into consideration site-specific factors. Program requirements and monitoring guidance are given in Environment Canada (1992; 1995a; 1995b; 1997; 1998; 2001). Key components of each EEM program include monitoring of fish, benthic invertebrates and supporting environmental parameters (e.g., effluent and water chemistry, sublethal toxicity testing and sediment chemistry) in reference and exposure areas to assess effluent effects. *This assessment is normally conducted using field survey techniques as they are the most environmentally relevant approach for effects assessment.* However, there may be situations where field surveys as recommended do not provide the required information. These situations could include confounded receiving environments and unsafe sampling conditions. If standard EEM field surveys cannot be conducted, use of scientifically

defensible, cost-effective, and technically feasible alternative monitoring methods can be considered.

Artificial stream systems or mesocosms are one of the recommended, scientifically defensible monitoring alternatives for EEM programs. See Section 11.2 for the information on the use of caged bivalves. Artificial streams are recommended as monitoring alternatives because years of research and development have demonstrated that they can produce good quality data on effluent effects that fit within the required regulatory context (Culp and Cash 1995; Culp and Podemski 1996; Culp et al., 1996; Culp and Lowell 1998; Culp et al., 2000a; 2000b; Dubé and MacLatchy 2000; Lowell et al., 2000; Culp et al., 2001; Dubé et al., 2001a; 2001b). Since 1991, the National Water Research Institute (NWRI) of Environment Canada has been developing field-based artificial stream systems for assessing the effects of point source effluents on aquatic ecosystems. Field-based artificial stream studies relevant to EEM applications have been conducted in Canada 14 times in 8 years from 1993 to 2000 (Table 1) (Dubé et al., 2001a). Table 2 summarizes the types of information that have been attained from these studies including the range of receiving environments studied, the types of effluents, specific endpoints for different levels of biological organization, and the types of hypotheses that have been tested.

In the Northern River Basins Study (1991-1996) a large artificial stream system was used as part of an integrated environmental assessment program to determine the effects of nutrients and contaminants from pulp mill effluent (PME) on benthic communities in the Athabasca River, AB (Culp and Podemski 1996; Culp et al. 1996; Podemski and Culp 1996; Culp et al. 2001). Similar studies were conducted in British Columbia from 1991-1998 in the Thompson and Fraser Rivers under the Fraser River Action Plan. In these studies artificial stream systems were used to improve the understanding of the effects of PME on freshwater riverine benthos (Culp and Cash 1995; Lowell et al. 1995; Culp et al. 1996; Dubé and Culp 1996; Lowell et al. 1996; Culp et al., 2000a; Lowell et al. 2000). In the late 1990's, artificial streams were used in three industrial pilot studies to assess the effects of PMEs on small-bodied fish in estuarine (Saint John River, NB) Monitoring (EEM) Programs (Dubé 2000b; Dubé et al., 2001a).

Table 11-1: Summary of artificial stream applications for assessing the effects of pulp and paper and mining effluents on aquatic ecosystems for Canadian Environmental Effects Monitoring

Year	Program[1]	Research Objective	Location	Artificial Stream System	References
1993-94	NRBS	To determine the effect of both nutrients and contaminants in PME[2] on periphytic algae and benthic invertebrate communities.	Athabasca River, Hinton, AB	Model I: Large non-mobile mesocosm	Podemski (1999) Culp and Podemski (1996) Podemski and Culp (1996) Culp et al. (1996) Culp et al. (2001)
1993-94	FRAP	To assess the effects of increasing concentrations of PME (0.25% -10%) on periphyton and chironomids (Diptera: Orthocladiinae).	Thompson River, Kamloops, BC	Small microcosm system	Dubé (1995) Dubé and Culp (1996) Dubé et al. (1997) Lowell et al. (2000)
1993	FRAP	To determine the food-dependent effects of PME on the mayfly (Ephemeroptera: *Baetis tricaudatus).*	Thompson River, Kamloops, BC	Small microcosm system	Lowell et al. (1995) Lowell et al. (1996) Culp et al. (1996) Lowell et al. (2000)
1994	FRAP	To determine the effects of PME (1% and 3%) on periphyton and benthic invertebrate communities.	Fraser River, Prince George, BC	Model II: Large mobile mesocosm	Culp and Cash (1995) Culp et al. (1996) Culp et al. (2000)
1997	Pulp & Paper EEM	To assess the effects of PME (3%) on a small-bodied fish (*Fundulus heteroclitus*) in a marine environment.	Saint John Harbour, Saint John, NB	Model II: Large mobile mesocosm	Dubé et al. (2001)
1997-98	Pulp & Paper EEM	To determine the effects of final PME, in-mill process streams, and a mill process change on a small-bodied fish (*Fundulus heteroclitus*) in an estuarine environment.	Saint John River, Saint John, NB	Model II: Large mobile mesocosm	Dubé (2000) Dubé and MacLatchy (2000a) Dubé et al. (2001)
1999	TSRI	To determine the effects of primary and secondary treated PME on a small-bodied fish (*Fundulus heteroclitus*) in an estuarine environment.	Miramichi River, Miramichi, NB	Model III: Large mobile mesocosm	Dubé et al. (2001)
1999-	TSRI	To determine the cumulative effects of PME and	Saint John River,	Medium modular	Culp, Dubé, Cash,

2000		municipal sewage on algae and benthic invertebrate communities.	Edmundston, NB	microcosm	MacLatchy, Glozier (unpublished)
2000	TSRI	To determine the effects of MME[3] on juvenile Atlantic salmon (*Salmo salar*).	Little River, Brunswick Mines, Miramichi, NB	Model III: Large mobile mesocosm	Culp, Dubé, Cash, MacLatchy, Glozier (unpublished)
2000	TSRI	To determine the effects of MME on algae and benthic invertebrate communities.	Little River, Brunswick Mines, Miramichi, NB	Medium modular microcosm	Culp, Dubé, Cash, MacLatchy, Glozier. (unpublished)
2000	TSRI	To determine the effects of MME on a food web including algae, benthic invertebrates and slimy sculpin (*Cottus cognatus)*.	Little River, Brunswick Mines, Miramichi, NB	Medium modular microcosm	Culp, Dubé, Cash, MacLatchy, Glozier (unpublished)

[1]NRBS: Northern River Basins Study; FRAP: Fraser River Action Plan; EEM: Canadian Pulp & Paper Environmental Effects Monitoring Program (EC 1992); TSRI: Toxic Substances Research Initiative. [2] PME: Pulp mill effluent. [3] MME: Metal mine effluent.

Table 11-2: Summary of types of information attained from Canadian artificial stream studies in the last decade (after Dubé et al 2001a).

Receiving Environment Type	Effluent Types[1]	Endpoints					Hypotheses Tested
		Primary Producers	Decomposers	Benthic Invertebrates	Community Metabolism	Fish	
• Freshwater • Estuarine • Marine	PME • BKM • TMP • Sulphite MME	• Biomass • Community composition • Indicator species • Tissue contaminants	• Decomposition rate • Fungal species composition • Fungal sporulation rate	• Abundance and biomass • Community composition • Emergence rates • Richness, diversity, ratio metrics • Developmental indicators (wingpad spread)	• Gross and net primary production • Microbial respiration • Oxygen isotope signatures	• Growth • Mortality • Condition • Liver size • Gonad size • Reproductive hormones • Tissue metals • MFO[2] induction • Contaminants in bile • "Stress responses" (muscle glycogen)	• Effluent effects • Cause of effect (in-mill waste streams, nutrients vs contaminants • Process changes • Cumulative effects

[1] PME = Pulp mill effluent; BKM = Bleached kraft mill; TMP = thermomechanical pulp mill; MME = Metal mining effluent. [2] MFO = mixed function oxygenase induction.

The most recent phase of artificial stream development has occurred since 1999 under the Toxic Substances Research Initiatives (Culp, Cash, Dubé, Glozier, MacLatchy, unpublished). In this program, stream systems have been developed to assess the effects of pulp mill and mining effluents on riverine and benthos, fish, and aquatic food webs (i.e., self sustaining multi-trophic level applications including algae + benthos + fish (Dubé et al. 2001a).

The objective of this document is to provide guidance on study design and implementation of mesocosm studies as an EEM monitoring alternative to assess the effects of PME and metal mining effluent (MME) on benthic invertebrates and fish. This Guidance Document is intended to provide information on recommended standards of good scientific practice available to meet the outlined EEM requirements. Please note that this Guidance Document is published in a format and is viewed as a document which can, and will, be updated regularly as new information and research becomes available. To ensure that the most recent updates are included in this document, consult our website at www.ec.gc.ca/eem, or contact Environment Canada's, National EEM Office (phone: 819-997-1535; fax 819-953-0461; e-mail: eem-esee@ec.gc.ca).

11.1.2 Applicability within EEM programs

It should be emphasized that mesocosms are a recommended monitoring alternative but their use should only be considered when field surveys cannot be conducted. Situations where artificial streams may be used include confounded receiving environments and where unsafe sampling conditions exist. Examples of confounded receiving environments include areas with the presence of historical effects, the absence of suitable reference areas for comparison to exposure areas, the presence of other effluents, and changes in relevant habitat types that cannot be factored out in the design of a field survey. There are advantages and disadvantages of using mesocosms and these should be weighed against the advantages and disadvantages of using other monitoring alternatives before final selection of an approach. Recommended alternative monitoring methods to the fish survey are mesocosm studies, caged bivalves, and expanded sublethal toxicity testing. For benthic invertebrate community surveys, mesocosm studies are the recommended alternative monitoring method.

11.1.2.1 Pulp and Paper EEM

Under the current *Pulp and Paper Effluent Regulations*, if field surveys cannot be conducted as outlined in the benthic invertebrate and fish survey decision trees then the use of alternatives is not a *requirement* until their use is included within the legislation (Environment Canada 1998). However, Environment Canada is currently developing proposed amendments to the *Pulp and Paper Effluent Regulations* and to the *Port Alberni Pulp and Paper Effluent Regulations* under the *Fisheries Act*. Included within these proposed amendments is the recommended use of

monitoring alternatives for effects assessment for fish and benthic invertebrates, including the use mesocosms. These regulatory amendments are proposed to come into force after the completion of Cycle 3 (April 1, 2004).

Although current EEM requirements are applicable until the regulation completes the amendment process at the completion of Cycle 3, mills that cannot conduct field surveys in Cycle 3 can propose the use of alternatives (e.g., mesocosms, caged bivalves) to their Authorization Officers. Mesocosms have been approved for use at some mills as alternatives in Cycle 2 and are proposed for use at other mills in Cycle 3. Thus, alternatives can be used in Cycle 3 if approved by the Authorization Officer even though their use is not *required* until legislated amendments to the *Pulp and Paper Effluent Regulations* are complete (start of Cycle 4).

11.1.2.2 Metal Mining EEM

11.1.2.2.1 Initial and Periodic Monitoring

Mesocosms can be used for assessing the effects of MME on benthic invertebrates and fish when standard EEM field surveys cannot be conducted (Environment Canada 2001). When effects are seen, it is essential that confirmation of these effects be seen in a subsequent monitoring phase (Periodic Monitoring - confirmation). However, at any time during the program, if an effect is seen, the mine can proceed directly to Focussed Monitoring or Investigation of Cause. If no effects are detected in a monitoring phase, it will also be necessary to confirm the absence of effects in a subsequent phase (Periodic Monitoring - surveillance). During the study design of the subsequent phase(s), it is crucial to evaluate whether the mesocosm design previously used, maximized the potential for detecting effects. Factors that need to be evaluated include species suitability, timing of the study, sample sizes, power of the analyses, and the level of effluent exposure in the mesocosm treatments.

11.1.2.2.2 Focussed Monitoring

When effects have been documented and confirmed through the completion of field surveys or mesocosm studies, it will be necessary to document the magnitude and geographic extent of the effect (Focussed Monitoring). The magnitude and geographic extent of effects can be assessed in mesocosms by measuring biotic responses to an ecologically relevant dilution series, by testing effluent effects on ecologically relevant or sensitive biota, or by testing effluent effects on ecologically relevant habitat types. One of the objectives of the subsequent phases may be to document whether other species demonstrate a similar effect (e.g., to assess the magnitude of the

effect). If desired, mesocosm study designs can provide flexibility to address Initial or Periodic Monitoring and Focussed Monitoring within the same study (Dubé et al., 2001a).

11.1.2.2.3 Investigation of Cause

Once the magnitude and geographic extent of the effect on fish is known, and previous monitoring has not determined the cause of the effect, identification of the cause of the effect is determined through Investigation of Cause. Mesocosms can be used for investigation of causal mechanisms associated with an effect. Lowell et al. (1995), Culp et al. (1996), Dubé and Culp (1996) and Culp et al. (2001) demonstrated how dose-response study designs could be used to separate the effects of nutrients from contaminants in PME on benthic food webs. In addition, Dubé (2000), Dubé and MacLatchy (2000; 2001), and Dubé et al. (2001b) showed how artificial streams can be used to isolate and identify the contribution of different in-mill process streams to final effluent quality and fish effects.

11.1.3 Mesocosm Technology

11.1.3.1 Suitability of Mesocosm Design

To maintain data quality and ensure consistent application of mesocosm technologies in EEM programs, review of the physical design of the system used, and the study design, including standard operating procedures, is required. This Guidance Document outlines licensing requirements, physical design guidance, and study design recommendations specific to a large transportable mesocosm system developed by NWRI.

Development of the large mesocosm system is on-going. A smaller modular design is also under development for use in EEM programs. As the technologies develop, design improvements and licensing opportunities will be described in amended Guidance Documents.

11.1.4 Description and General Operation of Large Transportable Mesocosm System

The mesocosm system recommended for use was developed by the NWRI and has been used to assess PME and MME effects on benthic food webs (benthic invertebrates and algae) (Culp and Podemski, 1996; Culp et al., 1996, Culp et al., 2000a; Culp et al., 2001; Dubé et al., 2001a) and small-bodied fish (Dubé 2000; Dubé and MacLatchy 2000; Dubé et al., 2001a; 2001b). The

stream system consists of 16 circular, 0.9 m^2 surface area tanks or streams placed in pairs on tables that are 74 cm high (Figures 1 and 2). Water from a reference area is pumped into a 378 L polyethylene head tank placed on a 1.2 m high platform, and gravity-fed through a system of pipes to the streams. Intake pumps such as a land-based pump (e.g., commercially available irrigation or pool pumps) or a submersible pump (0.5 HP Hydro-MaticJ Model # SPD50-H) can be used depending upon volume requirements, pumping distance, and the head required. Pumping reference water from a mill water intake is also a possibility. Gate valves control water flow to individual streams, and allow flow rate calibration for each stream. Water is delivered to each stream at a rate of 2 L/min resulting in a total water requirement of 32 L/min

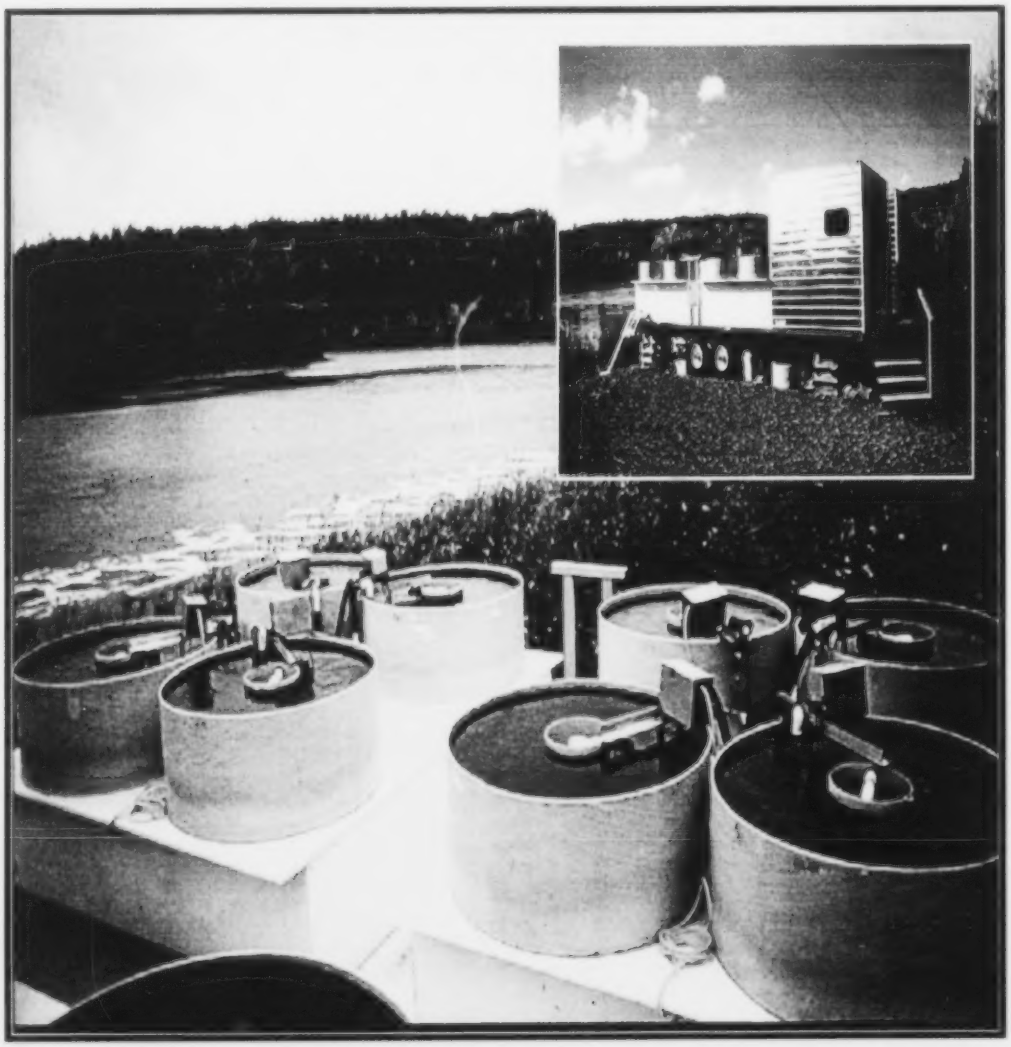


Figure 11-1: Picture of mesocosm trailer system used for EEM.

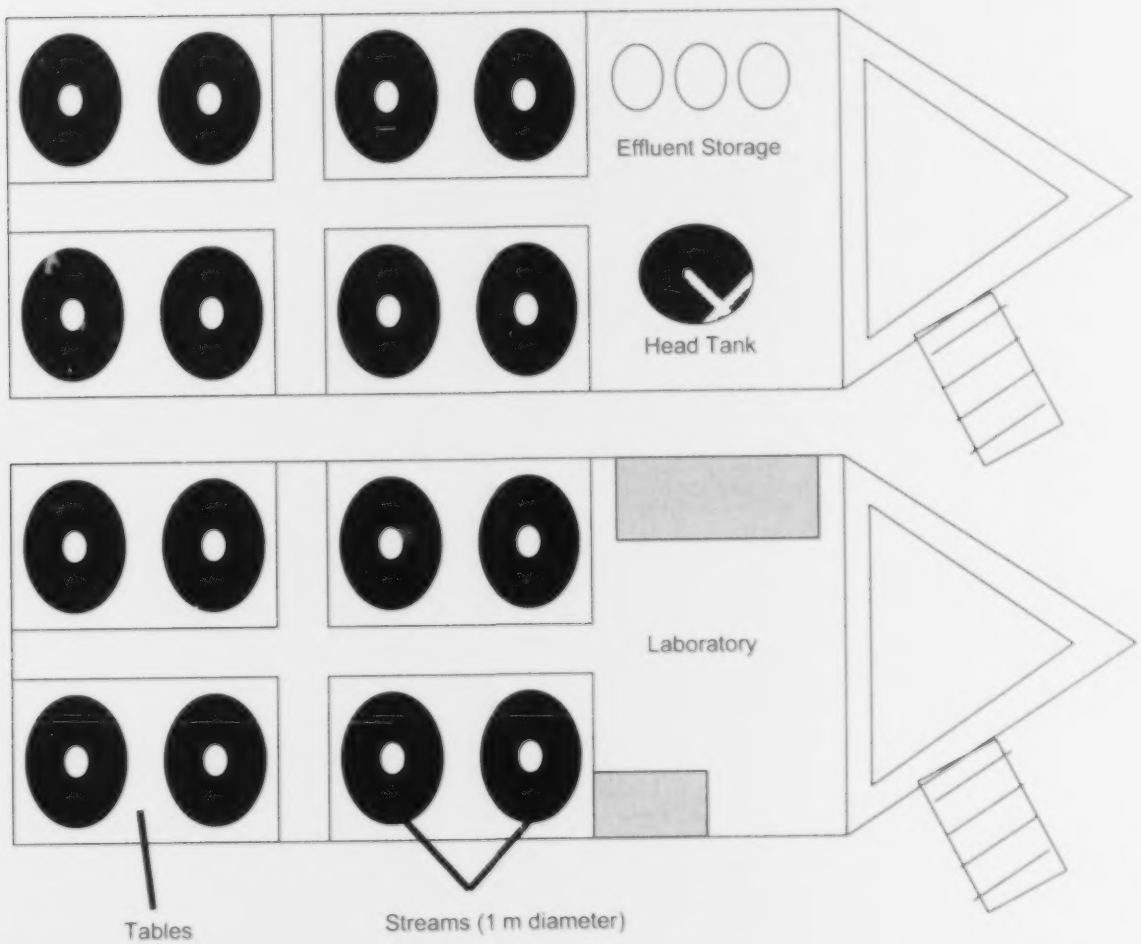


Figure 11-2: Schematic of mesocosm trailer system (not to scale)(after Dubé 2000)

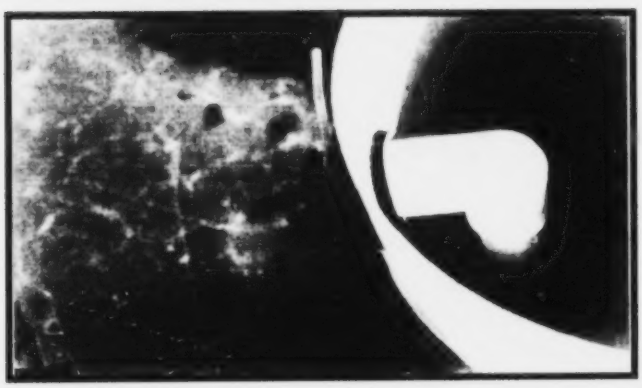
for the 16 stream system (Figure 3a). Water depth in the streams is maintained at 26.9 ± 0.1 cm ($\bar{x}$ ± 1 SE) by an overflow drain that returns all waste water to the river (Figure 3b). The overflow drains are screened to prevent fish loss and limit emigration of insects from the streams (Figure 3c). Each stream contains 227 L of water resulting in a hydraulic residence time of approximately 2 hours. By increasing water residence time within the streams (4 hours), the volume of effluents and reference water required during a study is minimized. However, we recommend an optimal stream residence time of ≤2 h. These recommended residence times are based on previous studies. However, final determination of residence time for a particular system depends upon the size of the tank used and species requirements for oxygen and temperature. If mesocosm studies are conducted in the late autumn, the head tank and water delivery lines can be wrapped with heat tape and insulated to allow the system to be operated when freezing temperatures may occur (-5°C). Shade cloth can be used to reduce solar heating in the summer.

The streams on the large mesocosm system are 107 cm diameter tanks constructed of polyester fiberglass (Figure 1). Streams are constructed by cutting 38 cm long sections of 107 cm pipe and bonding a flat sheet of fiberglass to one end of the pipe. A 25 cm diameter section of pipe is then centered in the larger pipe and the bottom cut out to form a standpipe. Streams are placed on eight, 74 cm high tables, two to a table. The water outflow pipe passes through the standpipe and drains into pipes beneath. These drain pipes connect to a general outflow pipe for discharge to the river downstream of the water intake point.

For invertebrate applications, current velocity in each stream is created by a belt-driven propeller system because this type of propulsive device is relatively gentle on drifting organisms (Figure 4) (Podemski 1999). The motor assembly for each stream is comprised of a geared head motor (250 RPM, 1/40 amp) driving a 22.6 mm pulley. A belt-drive transmits power from the motor to a 49.3 mm pulley mounted on the propeller shaft. The motor and associated electronics are mounted on an aluminum frame in a weatherproof enclosure. This frame is clamped to the top edge (outside) of the stream. A 16 x 230 mm long copper strut extending downward from the aluminum frame holds the propeller shaft bushing and grease seal. A grease nipple at the top of this tube allows for lubrication of the propeller shaft and bushing. The propeller (one per stream) is a 23 cm (9 inch) diameter aluminum fan blade which rotates at a no load speed of 115 RPM. Water velocity in each stream near the water column midpoint is normally maintained at 20 cm/sec although site-specific velocities should be the determining factor. Other current generation mechanisms can be used providing they produce the water velocities. Current for fish studies can also be attained using the propeller system, although spray bars attached to the water delivery system for each stream have also been used for species where water velocity is not a critical requirement.

Effluent is collected daily or every second day and stored in polyethylene containers. The point of effluent collection depends upon the study design. For an initial monitoring study at a pulp mill for example, the final effluent that is representative of what is being discharged to the aquatic environment is the target source. Effluent treatments are delivered independently and continuously to individual streams by peristaltic pumps

A



B

C

Figure 11-3: A) mesocosm stream
B) stream overflow drain on a mesocosm stream
C) screened overflow drain on a mesocosm stream

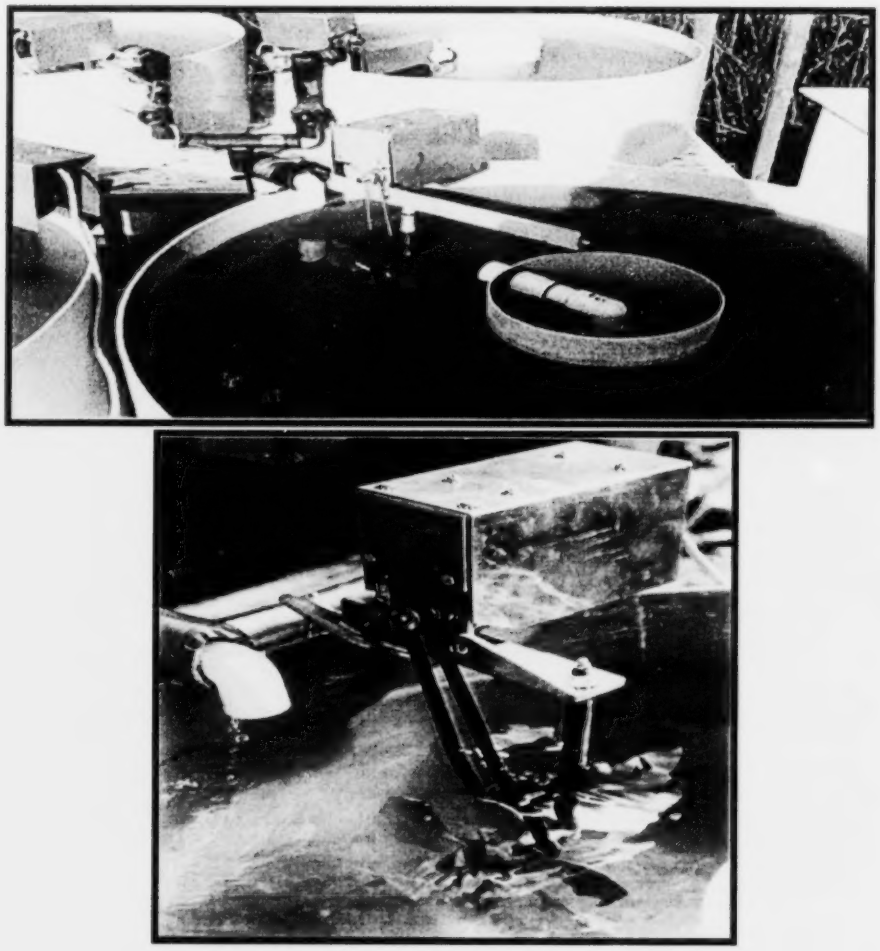


Figure 11-4: Fan propeller unit for current velocity.

(Masterflex ® L/S Nema-type 13 wash down controllers and cartridge pump heads) and a series of insulated tubes for solution delivery. Peristaltic pumps can be kept in insulated boxes to keep them within approved operating temperatures. Tubes carrying effluent can be threaded through foam pipe insulation to the streams, then fed into the water delivery spout on the water delivery pipe for each stream. Effluent flow rates depend upon the site-specific, environmentally relevant concentrations to be tested. For example, if plume delineation studies in the field have determined effluent concentrations to be 1%, then an effluent flow rate into each stream is set at 20 ml/min (1% of 2 L/min). If there are 16 tanks on the trailer, which are allocated between two treatments (8 streams for control; 8 streams for 1% exposure), then the following effluent volume is required on a daily basis: 20 ml/min X 60 min/hr X 24 hr/d = ~29 L/d/stream X 8 streams = ~231 L/d. This volume would fit into a small polyethylene container.

After water and effluent flows have been calibrated for each system, the individual tanks or streams are seeded with natural substrates, algae, benthic invertebrates, and/or fish species endemic to the receiving environment being studied. Inoculation of biotic populations is described below for fish and invertebrates. Further details on construction for this trailer system can be attained by contacting Dr. Joseph Culp (joseph.culp@ec.gc.ca) or Dr. Monique Dubé (monique.dube@ec.gc.ca) at the National Water Research Institute, Environment Canada, 11 Innovation Blvd., Saskatoon, SK, S7N 3H5.

11.1.5 Considerations for Site Selection

In order to operate the field-based mesocosm system there are some basic site requirements. The system is typically located at a reference area for access to reference water. In a freshwater riverine situation, reference water is pumped from upstream of the effluent discharge into the system and discharged back to the river downstream from the point of intake. This is the most straightforward scenario for mesocosm use. If the mesocosm is used in marine or estuarine applications, the site selected for the study must ensure that the quality of the reference water supply is not contaminated by effluent due to tidal influences. Effluent contamination includes both effluent discharged from the mill, or effluent discharged from the mesocosm system itself (i.e., outflow from the streams). In the former case, plume delineation studies are useful to determine the extent of effluent delineation in the receiving environment for appropriate selection of a reference site outside of the zone of effluent influence (Refer to Environment Canada 1998; 2001). In the latter case, the mesocosm can be located on a breakwater, land spit, or wharf to ensure the mesocosm outflow does not contaminate the inflow.

In mining applications, often the mine site is located in river headwaters and adequate reference areas can be a substantial distance from the mine site, particularly in remote areas. In addition, mining effluent concentrations are typically much higher than that found for pulp and paper discharge scenarios resulting in effluent requirements exceeding reference water requirements. Under these conditions, locating the mesocosm system at

the mine site and hauling and storing reference water is possible. In this scenario, the plumbing infrastructure on the trailer can be used for effluent delivery and the peristaltic pumps for reference water delivery.

Distance to the reference water source is also a consideration for site selection as extensive pumping requirements with respect to distance, or elevation (head) can exceed pump specifications for the rate required. Other site requirements include adequate space, site access, power and security. A 9 x 5 m area of level ground with good drainage is required for the trailer system. Access to power is one of the major site selection requirements. The system described in this document requires 120Amps of 115 Volt to operate the intake pump, the current generators, air pumps, etc. The mill typically provides this by installation of a power panel as an extension from an existing power line. In remote areas, if power is not available, generators have been used to power the system although this approach is not recommended. This is because maintenance and supervision of generators can make this approach cost prohibitive.

Once the site has been selected for the mesocosm study, the trailer is transported to the site, and water and power are hooked up. It is important that the trailer be properly oriented to minimize stream differences associated with light and wind exposure. One of the key benefits of mesocosms is that biota are exposed to effluents under environmentally relevant conditions of ambient temperature, water quality, and light regimes. If possible, it is recommended that the trailer be located in a north-south direction perpendicular to sunset/sunrise. The trailer should also be located away from the influences of shadows from buildings, etc that are in close proximity to the study area (i.e., all streams should receive the same duration and quality of sunlight).

11.1.6 Biological Monitoring Study Designs

11.1.6.1 Overview

The objective of a study design is to outline what mesocosm and associated laboratory work is needed to complete the biological monitoring portion of the EEM study. Many of the study components are similar to those outlined in the Technical Guidance documents for field surveys (Environment Canada 1998; 2001).

Prior to any biological monitoring studies being conducted, a study design will be submitted to the Authorization Officer at least six months prior to the commencement of conducting the biological monitoring program. The study design will include:

- a detailed timetable for conducting the mesocosm study;
- a description of the study area;
- a description of how the fish studies will be conducted in order to determine if there are effects on the fish population and fish tissue;

- a description of how the benthic invertebrate community studies will be conducted in order to determine if there are effects on the benthic invertebrate community as an indicator of effects on fish habitat; and,
- identification of the quality assurance and quality control measures that will be taken to ensure the validity of the data.

Other recommended details of the study design may include:

- defining the goals and expectations of the EEM study;
- determining the overall approach, including stating the rationale for choosing an alternative, which may be based on previous monitoring results;
- establishing statistical design criteria: development of hypotheses, selection of statistical methods, determination of data needs (statistical significance and power analysis);
- describing the design of the monitoring study: where to sample, what to measure, number of samples;
- developing operating plans and procedures: sampling procedures, laboratory analysis procedures, quality assurance and quality control procedures, data storage and retrieval, data analysis; and
- describing a plan for data interpretation and program evaluation.

The concepts and basic elements of a good measurement system, quality assurance/ quality control (QA/QC) are discussed in Environment Canada (1998; 2001). QA/QC procedures should be described in the study design and prior to the study. Quality assurance results that may affect the reliability of the conclusions will be submitted as part of each interpretative report.

11.1.7 Study Treatments

In mesocosm studies, the effect of effluent on fish is evaluated by comparing population indicators (growth, reproduction, condition and survival) between fish held under control conditions (reference water) to those held in effluent. For pulp and paper EEM programs, the effect may also be assessed by comparing levels of dioxins and furans in fish tissue between reference and exposure species. In mesocosm studies using benthic invertebrates, an effect is determined by comparing measurements of benthic invertebrate density, taxa richness, the Simpson's Diversity Index and the Bray-Curtis Index between invertebrates in reference versus exposure streams. It is crucial that mesocosm studies be designed to maximize the possibility of detecting effects if they are present. This includes using a *mesocosm design licensed for use by Environment Canada*; and selecting appropriate treatments, level of replication, sentinel species, endpoints, and conducting the studies at the proper time of year.

Several study designs can be employed depending upon site-specific requirements. The multiple tank mesocosm design provides flexibility for addressing site-specific considerations in the EEM study design with the 16 tank system resulting in up to three different treatments that could be assessed at the same time (five replicate tanks per treatment) under environmentally realistic exposure conditions. A control/impact design can be used where biota (i.e., benthic invertebrates or fish) held under reference conditions are simply compared to biota exposed to an environmentally relevant concentration of effluent (i.e., reference vs exposure). It is recommended that the environmentally relevant concentration of effluent is based upon a plume delineation study conducted during the EEM pre-design and represents the effluent concentration in the near field area after complete mixing (Environment Canada 1998; 2001). In the control-impact design for a fish program, typically adult (20 sexually mature males and 20 sexually mature females) and juvenile (20 sexually immature) fish of one small-bodied sentinel species are allocated to each stream (eight streams per treatment). Benthic invertebrates are stocked in the system at ambient densities found at the reference area in the dominant habitat.

In some cases, additional treatments may be desired. Dose-response study designs where an additional and higher effluent concentrations are used, is helpful to confirm biota responsiveness and the absence of effects. For example, Dubé and MacLatchy (2000) used three treatments (control vs 1% vs 5%) to assess the effects of a final effluent discharged from a bleached kraft pulp mill on a saltwater killifish (mummichog: *Fundulus heteroclitus*). In this example, effects were not observed at the environmentally relevant effluent concentration (1%) but did occur at 5%. This design proved invaluable to demonstrate that the fish were capable of responding to the effluent and the absence of effects at the 1% concentration was real and not a function of short-duration exposure studies. Similarly, Culp et al. (2000) used three treatments (control, 1%, 5%) to assess the effects of treated bleach kraft effluent on benthic food webs.

Additional treatments may also be desired if more than one mill discharges into the same sampling area. Each discharge can represent an exposure treatment in the experimental design and the effects of each effluent can be examined in isolation or in combination. For example, if a mill discharges three effluents into the same receiver then effluent effects can be evaluated using the following treatments: control, effluent 1 (environmentally relevant concentration), effluent 2, effluent 3, effluent 1+2+3 (environmentally relevant concentration). In this example, the number of streams used per treatment would decrease to three.

11.1.8 Replication

There has been considerable discussion surrounding the level of replication appropriate for mesocosm designs. For benthic invertebrate studies, individual streams are the unit of replication for the community-level endpoints. Thus in a control/impact design employing two treatments (reference vs exposure), n=8 as there are 16 tanks on the

mesocosm trailer. If additional treatments are preferred (see below), then 5 replicate streams per treatment should be the minimum number of replicates for EEM studies as this should provide adequate power for assessing effects when streams are the level of replication and when no previous monitoring data is available. This approach is consistent with the recommended method for determining the number of sampling stations, using statistical power in field survey designs (Environment Canada 1998; 2001).

The unit of replication is less clear when the system is used to measure individual-based endpoints in fish. The basis of the decision lies in quantification of the importance of a tank or stream effect (*i.e.,* an effect of one tank relative to another within the same treatment). In many laboratory studies where fish are held in aquaria, an assumption is made that there is no biological reason for tank differences within a treatment and thus all individual fish measurements are pooled within a treatment (*i.e.,* the variability attributable to tank effect is often ignored). For mesocosm studies we suggest that a tank effect may exist and the variation attributed to that effect requires some consideration in the statistical design. This is especially the case for longer-term exposures (e.g., 60 days). A nested ANOVA could be used in this example where the tank effect is nested within the fixed-effect factor of effluent treatment. If a tank effect exists then the level of replication should be at the stream level. If tank effects do not exist then fish can be selected as the unit of replication and pooled across streams within a treatment resulting in significant increases in the number of replicates. For regression-based analysis (ANCOVAs) any tank differences would likely fall out as outliers in the analysis. A description of the recommended statistical design for fish is outlined in Section 11.1.10.5.2.

11.1.9 Sample Sizes and the Role of Effect Size

11.1.9.1 No available pre-existing (historical) data

In mesocosm studies using benthic invertebrates, sample size represents the number of streams replicated per treatment as streams are the unit of replication. In a control-impact design, 8 streams are replicated per treatment. If additional treatments are required, then a minimum of five replicates or samples are required for each treatment to address power requirements without pre-existing data (see below). Each sample represents a mean of subsamples collected from each stream.

For a standard fish population field survey, where there is no background monitoring data, the minimum sample size recommended is 20 sexually mature males and 20 sexually mature females of two fish species collected from each of the reference and exposure areas. If small-bodied fish species are chosen as one or both of the fish species, an additional 20 sexually immature fish should also be collected. The rationale for using

20 fish of each sex is that there is little change in the 95% confidence limits with increasing sample size beyond 20 fish.

The current approach for mesocosm studies is to allocate 15 - 20 sexually mature males and 15-20 sexually mature females of one small-bodied sentinel species to each tank. In addition, 20-30 juvenile fish are allocated to the same tank. These are recommended sample sizes that can be increased or decreased depending upon the species selected, the statistical power of the study design, and fish variability. If tanks within a treatment are selected as the unit of replication then 20 fish per sex per tank provides a mean with good precision (i.e., there is little change in the 95% confidence limits with increasing sample size beyond 20 fish). If tank effects do not exist and error terms are pooled so that fish become the unit of replication, then 15-20 fish per sex per tank provides much higher replication per treatment. For example, a study design of control vs 1% effluent with 8 streams per treatment, and 20 males, 20 females, and 20 juveniles per tank can result in 160 males, 160 females, and 160 juveniles per treatment if fish are the unit of replication.

11.1.9.2 Available pre-existing (historical) data

If data are available, then the sample sizes to measure a certain effect size in a parameter with a targeted level of statistical power can be calculated because sample variability is known. Larger numbers of samples are needed if parameters are highly variable, if detection of small differences between reference and exposure streams (small effect size) is desired, or if a high level of power is required. To determine the sample sizes using pre-existing data, the effect size and the statistical power level that is acceptable for the decision-making process needs to be determined. The purpose of defining an effect size and power level is for determining when the sampling program is collecting sufficient information for decisions to be made. Sample sizes can be calculated using methods described in Environment Canada (2001). Appendix 1 of Environment Canada (1997) provides a detailed discussion of power relationships, effect sizes and the benefits of reducing variability in terms of increasing power. It is recommended that α and β be set equal at 0.10 as outlined in EC (2001) for power calculations.

The greatest factor affecting statistical power is sample variability. In general, mesocosm studies provide less sample variability compared to field studies as exposure is controlled and treatments are replicated. Mesocosm studies are designed site-specifically and priority should be given to reducing variability rather than increasing sample sizes. Variability can be reduced by randomly allocating streams to treatments, and biota (both benthic invertebrates and fish) to streams as described below in Sections 11.1.10.4.1 and 11.1.11.4.1. Random collection of samples at the end of the exposure period also helps to reduce variability to avoid collecting all samples from the same treatment at the same time (e.g., all control streams in Day 1 and exposure streams in Day 2). Other efforts to decrease variability include appropriate timing for the study, species selection, and QA/QC in sampling and analysis protocols. In addition, for fish studies, fish of uniform

size class are typically allocated to each stream to reduce variability. However, the use of small-bodied fish species, especially fractional spawners, may increase variability due to differences in spawning condition amongst individuals at the time of the study. Composite tissue samples for chemical analysis can also increase sample variability if the number of composites per stream is low.

11.1.10 Fish Monitoring –Effects Assessment

11.1.10.1 Study Design

11.1.10.1.1 Species Selection

The most important factors when selecting fish species for mesocosm studies are environmental relevance, abundance for collection, size, availability of adults and juveniles, sexual dimorphism, spawning period, and sensitivity to effluent. Local input may be valuable in providing information on the presence of fish species or information that may be helpful in determining sample locations for fish collection. The recommended method to assess effluent effects in mesocosm studies is by monitoring pre-spawning adults (sexually mature fish) and juveniles of one small-bodied fish species (e.g., darters, minnows, sculpins) that is relevant to the receiving environment. Monitoring of adults provides for assessment of effluent effects on survival, energy use (i.e., reproductive endpoints) and energy storage (i.e., condition, liver size). Monitoring of juveniles provides for assessment of effluent effects on survival, energy use (i.e., growth as a size-at-age substitute) and energy storage (i.e., liver size and condition).

The advantages and disadvantages of using small-bodied species in field surveys has been described by (Gibbons et al., 1998a; 1998b). A small bodied fish can be considered as a fish species that has a maximum size of 150 mm or less. Information on maximum growth of fish species can be found in the scientific literature, including Scott 1967; Scott and Crossman 1973; Fritz et al., 1975; Roberts 1988; Nelson and Paetz 1992; Jenkins and Burkhead 1993; Coad 1995; Leblanc and Couillard 1995; Leblanc et al., 1998. Shuter (1990) noted that small bodied fish species were more sensitive to acidification than larger species or at least responded more quickly to changes in pH. On a practical level, small bodied fish species are usually more abundant, easier to capture, and more sedentary than larger bodied fish species. In addition, measurements for the study can be made precisely and accurately on these small-bodied species.

There can be disadvantages to using small bodied fish. Often, less is known about their basic biology, particularly their spawning habits, making it difficult to determine the best sample areas, times and methods. Some species are also multiple spawners (produce more than one clutch of mature ova every year) (e.g., Heins and Rabito, 1986; Burt et al., 1988; Paine, 1990). This can be a disadvantage for field surveys because reproductive

effort in these species is difficult to estimate from a single sample. Reproductive tissue can be turned over almost completely between clutches (i.e., most of the mass of ova in the ovary will be spawned, and then a new clutch of mature ova will be developed). The number of clutches produced during the spawning season becomes the important reproductive variable and is difficult to estimate for an individual female in the field, even with frequent sampling. However, multiple spawners such as the mummichog (*Fundulus heteroclitus*) have been used successfully in mesocosm studies to evaluate effluent effects on EEM endpoints including changes in gonad size (Dubé and MacLatchy 2000; Dubé et al., 2001a; 2001b). An advantage of using controlled exposure mesocosm studies is that the state of fractional spawners can be monitored throughout the exposure period.

In addition to monitoring the effects of effluents on a single small-bodied fish species, the flexibility of the mesocosm system also allows for monitoring of two small-bodied fish species, or juveniles of large finfish. Theoretically, large-bodied finfish could theoretically also be used in mesocosm studies, although this approach has not been used previously. Species selection for mesocosm studies is restricted by size requirements (numbers of fish per tank decrease with increasing fish size), the number of treatments (increased numbers of treatments decreases the number of tanks available per treatment and limits the number of fish that can be monitored), and sexual dimorphism (if males and females cannot be externally sexed, then increased numbers should be considered to ensure adequate sample sizes). Sampled species should also be suitable for measuring the recommended variables. To date fish species used in mesocosm studies include mummichog (*Fundulus heteroclitus*), juvenile Atlantic salmon (*Salmo salar*), and slimy sculpin (*Cottus cognatus*). Fish proposed for use in 2001 and 2002 studies include creek chub (*Semotilus atromaculatus*), rainbow trout (*Salmo gairdneri*), longnose dace (*Rhinichthys cataractae*), and fathead minnow (*Pimephales promelas*).

11.1.10.2 Biological Endpoints

The endpoints measured in a mesocosm study using fish are the same as those measured in a field survey. To answer the question, "has the fish population been modified by effluent?", effects on growth, reproduction, condition and survival of the fish population are determined. Survival, energy use (growth and reproduction) and energy storage (condition, liver size) are measurements that can detect an effect in the fish population related to effluent. All of these measurements can be taken from fish exposed to effluents in mesocosm studies.

Growth is the change in size (weight or length) with time or age. Mesocosm studies over a longer exposure period (45-60 d) can assess effluent effects on juvenile growth, specifically changes in total body weight and length relative to control fish. In mesocosms, growth is measured in Y-O-Y or juvenile fish from the start of the study to the end; a duration of 6 to 8 weeks depending upon the temperature during the study. This is a more direct measure of growth effects than size-at-age estimates used in field

surveys. Determination of growth effects over a short exposure period requires that studies be conducted during warm water temperatures and with a sentinel species that has high growth rates. Kneib and Stiven (1978) have shown that juvenile growth rates of mummichog for example, are very high post-hatch (an increase of 15 mm in two months) which make them an ideal candidate for assessment of the effects of effluents on growth. More recent studies using Y-O-Y slimy sculpin have shown growth effects can be measured after only 26 days of effluent exposure (Culp, Cash, Dubé, Glozier, MacLatchy, unpublished). Growth measurements in juveniles can be improved further by tagging individual fish (fin clips, ink dots, branding) for length and weight measurements. Tracking the growth rate of specific individuals decreases the variability of these measures increasing the ability to assess effects. Growth can also me measured throughout the study at regular intervals if desired. Although length and weight measurements are taken in adults, growth is unlikely during the short duration of the mesocosm studies.

Reproduction is expressed as reproductive effort, fecundity, egg size or gonad weight relative to body size. To date mesocosm studies have examined changes in gonad size to assess effluent effects on reproductive function. However, measures of fecundity and egg size are easy to measure, if an appropriate sampling time is chosen. Ideally exposure studies should commence 6 to 8 weeks prior to the spawning season to assess effluent effects on gonad weight.

During the mesocosm study, the physical state of the fish is also assessed. Condition can be used as an index of well being and should be reported. A visual estimation of physical malformations including lesions on the body surface including: eroded, frayed or hemorrhagic fins; parasites; or other physical deformations, is recommended.

11.1.10.3 Timing and duration of effluent exposure

Timing of the mesocosm study is important and dependent upon the spawning cycle of the fish species. Ideally, for spring spawners, studies will be conducted 6 to 8 weeks before spawning commences. For early spring spawners for which it is impossible to study for this length of time prior to the spawning season, studies should be collected as late in the year as possible to allow for gonadal senescence and recrudescence. For fall spawners, a spring or summer survey is appropriate. This may not apply to fish in which ova mature rapidly. For example, some late-spring-spawning minnows should be studied in early spring, rather than in fall, when ova may still be immature. Water temperatures are also a consideration for timing of studies where juvenile growth rates are higher during the summer months. If historical data exists, it would be useful to examine the data and, if appropriate, conduct the study during similar periods so that the studies can be compared. Similarly, subsequent monitoring should be conducted during the same periods of the year to be comparable.

The duration of the mesocosm studies is a balance between being long enough to measure changes in the required endpoints and short enough to be cost-effective. Exposures of 30 d during summer and fall months have been used to measure changes in adult organ size (liver size, gonad size) and growth rates of Y-O-Y slimy sculpin as a result of effluent exposure. Longer exposures 50-60 days are recommended for assessment of both adult endpoints as well as changes in juvenile growth. Exposures beyond 60 days are not considered cost-effective unless cost reduction measures are taken (e.g., hiring a local summer student to maintain the system during the study).

All mesocosm studies should be conducted during normal industrial operations. Effluents are collected daily or every second day during the studies and should be representative of normal operating conditions. Storing of effluents during hardwood to softwood furnish shifts or during periods of bleach plant shut downs for example, should be avoided. Studies should be conducted when the mill is using a consistent furnish (i.e., half the study shouldn't be conducted during a softwood run and the other half during a hardwood run). Mesocosm studies at mills should not be conducted when effluent has not been discharged for long periods (months) or during waste water treatment upsets.

11.1.10.4 Fish Methods

11.1.10.4.1 General Set-up

Once the mesocosm system has been transported to the study site and reference water and power hooked up, then substrate is added to the streams, and water flow rates and stream current velocity are calibrated as previously described. Depending upon the site specific substrate characteristics, washed, crushed gravel or sand can be placed into each tank to a depth of 5 cm. In previous studies, large washed rocks were also added to each tank to serve as refugia for the fish.

Fish are collected from a reference area that is not exposed to the effluent being studied. Obviously, collection requires non-lethal sampling techniques including minnow traps, trap nets or electrofishing with barrier nets. These techniques are described in detail in the Technical Guidance Documents (Environment Canada 1995b; 1998; 2001). Fish are usually sorted in the field to ensure adequate numbers of juveniles, adequate numbers of adults of each sex (if sexually dimorphic) and to collect fish of similar size class. If the species chosen is not sexually dimorphic, an assumption is made that sex ratios in the field are equal and the fish should be randomly allocated to the tanks. It is possible to collect a subsample of these fish, dissect, and evaluate the sex distribution for additional information. Fish are transported back to the study site in containers with covers and under adequate aeration. If reference areas are not available, hatchery-reared fish can also be used if these fish are relevant to the study area.

At the study site, and preferably within the same day, fish are measured for length and weight, tagged if necessary and randomly assigned to each stream. The precision for these measurements is as described in the Technical Guidance Documents (Environment Canada 1995b; 1998; 2001) with increased precision when using small-bodied species. The optimal numbers of fish per stream are 20 adults of each sex and 30 juveniles. To ensure random allocation of fish to streams, 16 aerated buckets with reference water are set up. Fish are measured and placed into each bucket until each bucket contains one fish. This procedure is repeated until all buckets have five fish each. Buckets are then randomly allocated to the streams and the procedure is repeated until target sample sizes are attained in the streams. It is essential that this procedure be followed to ensure that the largest fish for example are not all allocated to the first few streams.

There are two possible fish allocation scenarios (Figure 5). Both adults and juveniles can be allocated to the same tanks if juvenile growth rates will be based solely on starting and ending length and weight measurements and adults are not cannibalistic (Figure 5a). If juveniles will be sampled during the course of the study for growth (perhaps juveniles are tagged and repeated measures are conducted) then allocating adults and juveniles to different tanks (or different areas within a tank if these areas can be physically separated) would be recommended to minimize adult capture stress (Figure 5b). Although Figure 5 indicates otherwise for ease of illustration, mesocosm streams are randomly assigned to each study treatment (control, effluent) and fish are randomly allocation to each stream. This ensures a randomized distribution to minimize stream differences due to factors such as stream position on the trailer.

After inoculation, fish are acclimated to artificial stream conditions for a minimum of 72 hours prior to effluent addition. Streams should be covered with netting to minimize fish loss due to escape or predation by birds. During the exposure period, which extends from 30 to 60 d, fish should be fed trout pellets at a rate of 4-6% total body weight per tank per day. The size of the pellets is dependent upon the size of the species used. In previous studies, trout pellets were crushed prior to feeding.

11.1.10.4.2 Maintenance and Monitoring

Daily maintenance for the duration of the study includes calibration of both water and effluent flows to each stream to ensure target effluent dilutions are attained; cleaning of screens; and feeding of fish. Screens are placed over the water outlet pipes on each stream to prevent fish loss (Figure 3c). These screens require daily cleaning to keep the overflow drains clear. Pump maintenance and changing of peristaltic lines due to clogging, algal growth, or leaking due to frictional wear over the pump roller heads, is also monitored and maintained on a daily basis.

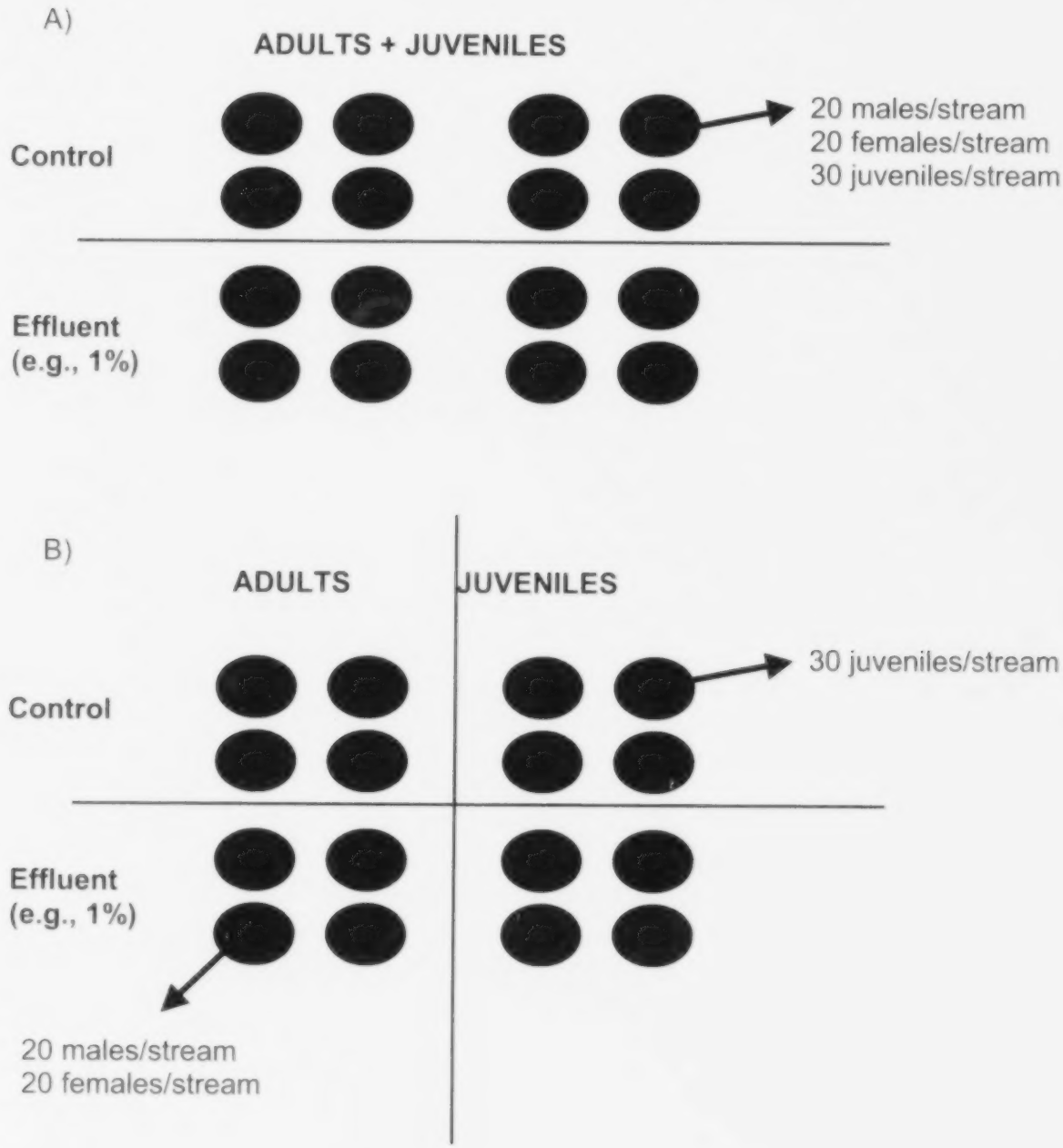


Figure 11-5: Fish allocation scenarios where A) adults and juveniles are allocated to the same stream or B) adults and juveniles are allocated to separate streams. For actual implementation, streams are randomly allocated to each treatment.

Daily monitoring requirements for each stream include recording of any fish mortalities and monitoring of physical water chemistry including temperature, dissolved oxygen, conductivity, and pH. These measurements can be taken using YSI instruments or continuous recording equipment such as hydrolabs or thermisters. Dissolved oxygen is of critical importance especially in studies using effluents high in organic content (e.g., pulp mill effluents). Dissolved oxygen levels should be maintained above 60% (minimum) and preferable above 80% in all streams, using aeration if required.

11.1.10.4.3 Sampling and Analysis of Fish

The order of stream sampling should be random and preferably all streams should be sampled on the same day. If this is not possible, random sampling of the streams limits differences amongst streams within a treatment due to differences in the time of sampling over two days. All fish from each stream should be sampled before progressing to the next stream.

To sample a stream, water and effluent flows to that stream are stopped and water depth is reduced using screened siphons. Large rocks can be removed to improve ease of sampling and fish are collected using hand nets. It is easiest if males, females and juveniles are sorted into three separate buckets at the time of collection. Fish are held under cover and aeration until measurements are taken. An efficient 4-person measurement team should be able to process all fish from one stream in half an hour on the average.

The measurements required and level of precision are the same as those outlined in the Technical Guidance Documents (Environment Canada 1998; 2001). Sample preparation, laboratory analysis and QA/QC procedures for mesocosm studies are the same as those required for field fish surveys.

11.1.10.5 Data Assessment and Interpretation

11.1.10.5.1 General Requirements

For mesocosm studies, Data Assessment and Interpretation follows each monitoring or assessment phase in a similar way as for field surveys. In Data Assessment and Interpretation, the following questions are answered:

- is there an effect?
- is the effect mill related?
- is the magnitude and extent of the effect known? and,
- is the mill related cause of the effect known?

For the purposes of the EEM program, an effect on fish is defined as "a statistical difference between fish population measurements taken in an exposure area and reference area". An overview of data analysis and interpretation for mesocosm studies is presented here. However, Environment Canada (1998; 2001) should be consulted for further details.

Following Data Assessment and Interpretation, the mills will submit an interpretative report to the Authorization Officer (AO). This report will include information such as description of any changes from the submitted study design, the identification of any effects, the conclusions of the study, and a timeline outlining when the next monitoring phase will be conducted.

11.1.10.5.2 Statistical Analysis of Parameters

To determine whether or not there is an effect on fish exposed to effluents in mesocosm streams, statistical analysis of the data are conducted as suggested in Tables 3 and 4. The biological variables will be used to derive the following parameters:

Energy Use (Growth and Reproduction)

- growth: size of juveniles at end of the exposure period compared to starting measurements
- reproductive investment: gonad weight (as gonad weight relative to body weight; gonad weight relative to length), fecundity (as fecundity relative to body weight; fecundity relative to length)

Survival

- percentage surviving in the streams at the end of the exposure period

Energy Storage (Condition)

- condition (body weight relative to length)
- liver weight (as liver weight relative to body weight and liver weight relative to length).

Sex differences in growth rate, body weight, condition factor, gonadal size and liver size are common due to differences in overall energetic requirements between male and female fish. Therefore, for all parameters, sexes should be treated separately when estimating variability. Immature or juvenile fish should also be treated separately. The analyses that are suggested in Tables 3 and 4 are preferred. However, other analyses for specific parameters can been conducted depending upon the variability of the data set. For example, in mesocosm studies individuals of a similar size class are selected for placement in the streams resulting in reduced variability in endpoints compared to those measured in field surveys. Thus the range over which regressions such as ANCOVAs are conducted for mesocosm data is small. ANOVAs on original organ weights (liver size, gonad size), and preferably ratio metrics (LSI, GSI), or condition factors are justified in this instance.

The order of data analysis is largely a matter of personal preference, but the recommended methods are divided into five stages: a) preparing the analyses, b) initial summary statistics, c) ANOVA analyses, d) ANCOVA analyses, and e) power analyses. Details for these stages are available in Environment Canada (1998; 2001). Statistical considerations specific to mesocosm data sets are stated below.

Table 11-3: Effects on fish. Biological endpoints recommended for determining effects as designated by statistically significant differences between exposure and reference streams.

Type of Response	Effect Endpoint	Statistical Procedure
Survival	Juvenile and Adult Survival	ANOVA
Energy Use (Growth)	Juvenile Growth (change in length over time)	ANOVA (nested)
	Juvenile Growth (change whole animal wet weight over time)	ANOVA (nested)
(Reproduction)	Relative gonad size (gonad weight against body weight) - Adults	ANCOVA
Energy Storage (Condition)	Juvenile and Adult Condition (body weight against length)	ANCOVA
	Relative liver size (liver weight against body weight)- Adults	ANCOVA

Table 11-4: Biological endpoints to be used for supporting analyses. These analyses are for informational purposes and significant differences between exposure and reference streams are not necessarily used to designate an effect.

Type of Response	Supporting Response Variable	Statistical Procedure
Energy Use (Growth)	Adult Body weight (whole)	ANOVA (nested)
	Adult Length	ANOVA (nested)
(Reproduction)	Relative gonad size (gonad weight against length)- Adults	ANCOVA
	Relative fecundity (# of eggs/female against body weight)	ANCOVA
	Relative fecundity (# of eggs/female against length)	ANCOVA
Energy Storage (Condition)	Relative liver size (liver weight against length)	ANCOVA
	Relative egg size (mean egg weight against body weight)	ANCOVA

11.1.10.5.3 Nested ANOVA Analyses

For any variable measured once on a whole replicate stream, the statistical design is a simple t-test, comparing exposed and control treatments (control/impact design). The mesocosms or streams are replicates for the treatments. However, all biological and biochemical endpoints are measured on replicates at a lower level - either individual fish or composites of several fish within streams. A nested ANOVA is used to analyze these variables (Table 5). If the variance among replicate streams within a treatment (Error I) is large relative to the variance among fish or composite samples within streams (Error II), then the nested ANOVA is effectively a t-test comparing treatments, with the stream means as replicate observations. However, if Error I is not large relative to Error II (e.g. $p>O.25$), then the two error terms can be pooled to increase the power of the test comparing treatments (i.e. individual fish or composite samples rather than streams are used as replicates).

Table 11-5: ANOVA Table for Nested Design, with Two treatments, Four replicate Streams per treatment, and *n* Fish or Composite Samples Within Each Replicate Stream

Source		Df	F-test
Treatment	Tr	(Tr-1) = 1	MS (Tr) /MS (Error 1)
Streams within treatments	Error 1	Tr(r-1) = 6	MS (Error 1) /MS (Error 2)
Fish or composite samples within streams	Error 2	T(n-1) = 232	

Where: df = degrees of freedom; MS Mean Square
Tr = number of treatments, r = number of replicate tanks per treatments
T = number of tanks, n = number of fish per tank

11.1.10.5.4 ANCOVA Analyses

Most parameters are estimated as regressions of one variable on another using Analysis of Covariance (ANCOVA). ANCOVA combines the features of analysis of variance and regression, and can be used to compare regressions among multiple classifications (i.e. sampling areas) (Snedecor and Cochran 1980). The mean squared error (MSE; model variance, $s^2 = SD^2$ of the residuals from the regression]) of the ANCOVA model should be used to provide an estimate of variance of the fish parameters, when reasonably large samples (e.g., including 20 or more fish) are used. However, in mesocosm studies, the fish are chosen so that they are all within a comparatively narrow size/age range. Thus, it may not be necessary to use ANCOVA to factor out the influence of the size/age covariate on the endpoints. If limited fish availability necessitated the use of a wider size/age range, then ANCOVA comparisons might be required for some of the endpoints to factor our size/age effects.

11.1.10.6 Data QA/QC and Analysis

The importance of ensuring data quality can not be overstated. There are basics related to study design, consistency of methods and measurements, definition of protocols and procedures, which are outlined in detail in Environment Canada (1998; 2001).

Data should be tested for validity of statistical assumptions including normality and homogeneity of variance before applying parametric procedure as described in Environment Canada (1998; 2001). However, most univariate normal distribution-based statistical based methods are quite robust and can support moderate violations of the assumptions. Transformation of original data will help normalize the data or homogenize

the variances. Logarithmic transformation are often preferred because most biological measures are considered to operate on a log or exponential scale (Peters, 1983). If the transformations are unable to produce data that meet the assumptions, then non-parametric statistics are used.

When designing experiments, it is also desirable to ensure that replicates are randomly allocated to different treatment levels such that the responses of each replicate is independent of other replicates. This element of randomness provides some assurance that observed differences in responses among treatments results from treatment-effects and not for some other reason. Randomly allocating replicates to different treatment levels is a relatively easy procedure when conducting manipulative experiments (e.g., mesocosm studies), but is less obvious for observational field studies. In natural environments, the high degree of spatial heterogeneity and the challenge of obtaining true replicates makes it difficult for biomonitoring studies to predict quantitatively the impacts of complex effluents on biota (Buikema and Voshell 1993, Forbes and Forbes 1994). Although biomonitoring provides important observations and insights into patterns of effluent effects on biota, inferential hypothesis testing through field assessments is often impossible (Stewart-Oaten et al. 1992; Cooper and Barmuta 1993). Experimentation using artificial streams is desirable in that they provide control over relevant environmental variables and true replication of treatments (McIntire 1993, Culp and Podemski 1996). Thus, causality (i.e., cause and effect) can be assigned definitively.

11.1.11 Benthic Invertebrate Community Monitoring Effects Assessment

11.1.11.1 Study Design

11.1.11.1.1 Species Selection

The benthic community that is established in the mesocosm streams is representative of that found in the reference field areas. Benthic samples are collected from the river and the entire community is inoculated into each mesocosm stream. Culp et al. (1996, 2001), Podemski (1999) and Culp et al. (2000a) have conducted studies that compared the community structure of benthic invertebrates in the mesocosm streams to that of field communities at reference areas. No ecologically significant differences in structure were observed, illustrating the effectiveness of the inoculation procedures and the suitability of mesocosms for testing effluent effects on environmentally relevant communities of benthic invertebrates.

The effectiveness of artificial streams in assessing effluent effects on environmentally relevant riverine communities was further corroborated in mesocosm studies on the Athabasca River (Culp and Podemski 1996; Culp et al. 2001). In these studies, effluent

exposure significantly increased algal biomass and benthic insect abundances in the mesocosm streams. These results corresponded with in-river trends where periphyton biomass and benthic insect abundances increased at field sites downstream of the effluent (exposure areas) relative to upstream reference areas. These artificial stream studies substantiated speculation from earlier field studies which attributed higher abundances downstream of the outfall to pulp mill effluent-induced nutrient enrichment (Anderson 1989, Anderson 1991). Improved understanding of the mechanisms by which pulp mill effluent exposure affects the river ecosystem was achieved because the use of the stream-side mesocosms allowed the separation of confounded environmental factors at a community level (Culp and Podemski 1996).

11.1.11.2 Endpoints

11.1.11.2.1 Biological Endpoints

The objective of the benthic invertebrate component of the EEM programs is to determine if the mill effluent has modified the benthic invertebrate community (i.e., is there an effect?). The mesocosm system can be used to assess the effects of effluents on a variety of benthic endpoints including those of community structure and process, as well as the growth and fecundity of selected community members (Culp et al. 1996, Culp and Podemski 1996, Podemski and Culp 1996, Culp et al. 2000, Culp et al. 2001). Community-level studies incorporate greater environmental realism (i.e., are more accurate in terms of incorporating the modifying effects of interspecies interactions on overall stressor effects), are better able to identify those components within the community most sensitive to a given effluent, and integrate a variety of ecological processes (e.g., competition and food web dynamics) that cannot be simulated in single species studies (Forbes and Forbes 1994). Furthermore, multispecies tests are a very efficient way to obtain information on the relative sensitivity of a large number of species in a single study. In mesocosm studies, we have been able to address community-level questions relating to the potential effects of varying concentrations of effluent on primary (attached algae) and secondary (benthic insects) producers. Specific endpoints measured in these studies included periphyton growth, the growth of selected insect taxa, algal and invertebrate community structure, and contaminant concentrations in algae (Table 1) (Culp and Cash 1995, Culp et al. 2000, Culp et al. 2001, Dubé et al., 2001a).

All endpoints used to assess effluent effects on benthic invertebrates in EEM field surveys can be measured in mesocosm studies. For EEM, an effect on the benthic invertebrate community is defined as a statistical difference between benthic invertebrate community measurements taken in an exposure area and a reference area, or a statistical difference between measurements taken at sampling areas in the exposure area that indicate gradually decreasing effluent concentrations, for any of the following key descriptors:

- Total invertebrate density
- Taxon (i.e., Family) Richness
- Simpson's Diversity Index
- Bray-Curtis Index (summarizes the overall difference in community structure between reference and exposed treatments in a single number and requires no preconceived assumptions about the nature of the community (Taylor and Bailey 1997)).

Evenness, taxon density, proportion or presence/absence are recommended to allow for the interpretation of effects (Environment Canada 2001). These descriptors are summary metrics selected to encompass the range of responses which may result from effluent, including changes in productivity, species composition and biodiversity. Many other benthic invertebrate descriptive metrics are available in the literature (for a review see Resh and McElravy 1993) and may be used, if applicable, on a site-specific basis to aid in the interpretation of effects determined with the key descriptors listed above.

11.1.11.3 Timing and duration of effluent exposure

Mesocosm studies using benthic invertebrates can be conducted at any time from early spring to late fall. The primary limiting factor is temperature with temperatures below -5 °C prohibitive to mesocosm operation due to freezing of water lines. Studies should be conducted during periods when field communities are under maximal effluent exposure for improved environmental realism. Studies should also be conducted at the time of year when the benthic invertebrate diversity is highest and water temperatures are conducive to growth. If historical data exists, it would be useful to examine the data and, if appropriate, conduct the study during similar periods so that the studies can be compared. Similarly, subsequent monitoring should be conducted during similar periods of the year to be comparable.

The duration of the mesocosm studies for benthic invertebrate programs is typically 30-45 days which includes a 7-12 day inoculation period for primary producers and a 25-30 day effluent exposure period. All community endpoints can, and have, been measured over this study duration. Effluents are collected daily or every second day during the studies and should be representative of normal industrial operating conditions.

11.1.11.4 Benthic Invertebrate Methods

11.1.11.4.1 General Set-up

General set-up requirements specified in Sections 11.1.3 and 11.1.6 also apply for benthic invertebrate studies. The recommended total flow of reference water or of dilution water and effluent into each stream is 2 L/min. This flow rate results in a 2 hour volume turnover time for each stream (stream volume 227 litres). Stream volume turnover rates of 4 hours have been used to reduce effluent requirements. However, 4 hour turnover times are the maximum. Longer turnover rates are not recommended due to effects on stream temperature and dissolved oxygen. For invertebrate applications, current velocity in each stream is created by a belt-driven propeller system as described in Section 11.1.3. Current in each stream is established based on site-specific conditions but is normally set between 10-20 cm/sec.

Prior to placement of substrate into the streams, five sampling bags (0.1 m^2, 500 µm mesh) are installed on the stream bottom. These bags are lifted at the end of the experiment resulting in five subsamples per stream. A standardized benthic environment is created in each stream to simulate the dominant, environmentally relevant habitat type. To date, mesocosm studies using benthic invertebrates have focused on riverine habitats where riffle substrate is dominant. The bottom of each stream and the sampling bags are covered with an 8 cm layer of washed gravel (1-2 cm diameter stones). The gravel is then left to colonize for a 7 to 12 day period to allow sufficient algal and microbial growth. Only water delivery to the mesocosm streams occurs (2 L/min) during this colonization period. The duration of the colonization period is site-specific and depends upon temperature and colonization of algae from the river into the system.

Following the algal colonization period, existing benthic communities are transplanted to the stream mesocosms. This protocol is determined on a site-specific basis. The following is an example of inoculation protocol for a riverine, riffle habitat with a large cobble/gravel dominated substratum:

1. Large cobbles (~x surface area = 535 cm^2) are randomly selected from the river at the reference area for placement into the mesocosm streams with their associated periphyton and invertebrate biota. These cobbles provide additional substrate and also stock the streams with a natural community of periphyton and benthic invertebrates.
2. During collection, the stones are enclosed with a 0.1 m^2 U-net (500 µm mesh) (Scrimgeour et al. 1993), carefully lifted from the streambed and placed into a container with river water. In addition, the gravel substratum beneath the stones is gently disturbed (to a depth of 5 cm) to collect any invertebrates under and around the base of the stone. The container is carefully transported to the mesocosm so that the periphyton and invertebrates associated with the stone are not dislodged. The cobbles are randomly placed in the artificial streams orientated to the water

flow similar to their original orientation in the natural environment. This process continues until the appropriate density (e.g., 10 large cobbles/artificial stream) of large cobbles is reached based on natural substratum composition. Other samplers may be used to inoculate the streams with benthic communities as outlined in Environment Canada (1998; 2001). The objective however, is to establish ambient densities in the mesocosm streams. To determine the number of river samples to place in the streams, the mesocosm stream area (0.9 m^2) is divided by the area of the selected sampler. This random allocation of subsamples from pooled invertebrate community samples ensures that the initial invertebrate composition is similar amongst streams and limits the amount of variability in community composition among mesocosms attributable to the pattern of species introduction. (Wrona et al. 1982)

3. An additional series of benthic samples (biota only) are collected from the reference area and are pooled into a common container to estimate initial composition and density. Sub-samples (Marchant 1989) are then removed and randomly inoculated into each artificial stream until the density of invertebrates approximates ambient densities.

Standardization of inoculation protocols is important because the sequence through which species from a common source pool are added to mesocosms can produce large differences in community structure; these dissimilarities are unrelated to the intended study treatments (Drake et al. 1996). The test community includes multiple trophic levels constructed from random samples from the study river. Consequently, measurements of community-level endpoints, such as species composition, community production/respiration, and decomposition, are possible. These community-level measures integrate both the direct effects of stressors and, importantly, indirect ecological effects that cannot be simulated in single species or simple food chain systems (e.g., competition-mediated shifts in community structure, or biomagnification) (Carlisle 2000).

Following biotic inoculation, the mesocosm system is left to acclimatize for 24-48 hours (water flow only) prior to commencement of effluent delivery to the treatment streams.

11.1.11.4.2 Maintenance and Monitoring

Maintenance and monitoring is as described for the fish in section 11.1.10.4.2. Streams are inspected daily and drain screens are brushed clean to prevent clogging and overflow of the streams. Motors are lubricated with non-toxic, Permatex Superlube ® (food grade USDA H1) every 3-6 d to prevent seizing of the propeller shafts. The water delivery lines are flushed regularly to remove silt and sand deposits. Water and treatment solution flows to each stream are calibrated regularly. Effluent delivery tubing also requires daily inspection for blockages and wear from the pump heads, and changed as required.

Daily monitoring requirements for each stream includes monitoring of physical water chemistry including temperature, dissolved oxygen, conductivity, and pH. These measurements can be taken using YSI instruments or continuous recording equipment such as hydrolabs or thermisters. Dissolved oxygen is of critical importance especially in studies using effluents high in organic content (e.g., pulp mill effluents). Dissolved oxygen levels should be maintained 60% (80% preferred) in all streams using aeration if required.

11.1.11.5 Sampling and Analysis of Algae and Benthic Invertebrates

Sampling of algal communities in the mesocosm streams is considered a supporting measure to assessing effects of effluents on benthic invertebrate communities. However, the information for sampling and collection is presented here as these samples must be collected prior to sampling of benthic invertebrates to ensure sample integrity. Sampling of streams is done through randomized selection to minimize differences due to the time of sampling amongst replicate streams within the same treatment. Periphyton samples are collected once at the end of the exposure period, and a composite sample of three rocks per stream are sampled in order to reduce variability. Rocks are sampled by using a scalpel to remove periphyton from within a template of known area (e.g., 9.6 cm^2). Periphyton samples are placed in vials and held on ice until frozen later the same day. In the laboratory, each sample is homogenized, partitioned into two parts, and each portion filtered through a GF/C filter. Chlorophyll *a* concentration is determined by extracting the filter and retained material in an 80 °C bath of 90% ethanol for 5 min, then measuring fluorescence with a Turner Designs, model 10 series fluorometer. Ash-free dry mass (AFDM) is determined by weighing the sample after drying for 24 h at 105 °C, then combusting the filter at 500 °C and determining the weight loss upon ignition.

For benthic invertebrate sampling, streams are also sampled in a random selection to minimize differences due to the time of sampling amongst replicate streams within the same treatment. Invertebrate samples are collected by lifting the sampling bags within each stream which results in five subsamples per stream. Subsamples are then washed through a 500 µm mesh sieve. Field sieving is required immediately after sample retrieval and before preservation as described by Environment Canada (1998; 2001). The recommendation for sieve and/or mesh size for all freshwater mesocosm applications is 500 m. In freshwater, macroinvertebrates are defined as those retained by mesh sizes of 200-500 µm (Slack et al. 1973; Weber 1973; Wiederholm 1980; Suess 1982), although immature life stages of some taxa may be smaller and some adult life stages may be larger. Note that these mesh sizes are applicable to all equipment used in the field and laboratory (i.e., both the Nitex mesh on the benthic samplers and sieving apparatus). In some site-specific circumstances it may be desirable for the samples to be screened for smaller organisms by using a smaller sieve size. Some examples of situations where the use of a smaller mesh size (less than 500 m) may be appropriate include the following:

1) for comparative purposes if historical benthic surveys for the system under investigation utilized smaller mesh sizes, or
2) if sampling needs to be conducted, for logistic reasons, at times when organisms are very small Rees (1984), Barber and Kevern (1974) and Jonasson (1955) provide information on seasonal effects of mesh size.

In these aforementioned cases, it is highly recommended that a stack of screens be used which minimally have the mandatory sieve sizes and then any other smaller sizes which are appropriate. This procedure simultaneously allows site-specific concerns to be addressed and fulfil the EEM objectives by allowing for national or regional comparisons to be conducted on the standardized mesh sizes.

After sample sieving, invertebrates are preserved immediately in 10% formalin. In the laboratory, samples are sorted and identified. In previous mesocosm studies, the contents of three randomly chosen sampling bags were sorted from each stream. The recommended level of taxonomic identification is *family*. All summary statistics and descriptive metrics should be calculated and reported at the family level for submission for EEM interpretative reports. Organisms that cannot be identified to the desired level of taxonomic precision should be reported as a separate category in the fundamental data set. It is recommended that investigators use taxonomic keys appropriate to the geographic region of study; a detailed list of taxonomic references for various groups of freshwater organisms is provided in Environment Canada (1998; 2001)

11.1.11.6 Data Assessment and Interpretation

Use of mesocosms as a monitoring alternative for assessing effluent effects on the benthic invertebrate community largely follows the guidance of Environment Canada (1998; 2001) for data assessment and interpretation in field surveys as the endpoints measured to assess effects are the same. The only difference in data assessment is replication per treatment in mesocosm studies is n=8 for control/impact designs rather than n=5 as recommended for field surveys. Data Assessment and Interpretation follows each monitoring phase with the objective of answering the following questions:

- is there an effect?
- is the effect mill related?
- is the magnitude and extent of the effect known? and,
- is the mill related cause of the effect known?

During the effects assessment, a significant difference between reference and exposure areas in any of the following key descriptors is to be interpreted as an effect on the benthic invertebrate community.

1) Total invertebrate density
2) Taxon (i.e., Family) Richness

3) Simpson's Diversity Index
4) Bray-Curtis Index

Evenness, taxon density, proportion of presence/absence are recommended to aid in the interpretation of effects (Environment Canada 2001). Calculation of these metrics is described in detail in Environment Canada (1998; 2001). Details on recommended statistical analyses, data QA/QC, and reporting requirements is as outlined for field surveys with the qualification that a mesocosm treatment level is equivalent to an exposure area and replicate artificial stream is equivalent to a field replicate station.

11.1.12 Supporting Measurements

11.1.12.1 Water Chemistry Endpoints

In mesocosm studies, it is essential that stream differences in physical water chemistry be minimized as much possible to ensure effluent-related effects are not confounded. Streams are monitored daily for temperature, dissolved oxygen, pH, and input flow for water and effluent. Current velocity should be measured at the start and end of each study, especially for benthic invertebrate studies. The distribution of water velocities in the streams is characterized using one of many brands of a velocity meter. In previous studies, mean velocity in the mesocosms (above stones) ($x = 0.26 \pm 0.01$ m/sec, $n = 150$), was similar to water velocity measured above stones at a similar water depth in the field (water velocity $x = 0.26 \pm 0.01$ m/sec, water depth $x = 24.8 \pm 0.72$ cm, $n = 30$) while the study was being conducted (Podemski 1999).

Water temperature can be measured using continuous temperature loggers. In the latter instance, water temperature can be monitored with data loggers placed in a stream and the head tank. Temperatures in the head tank reflect the temperature of incoming river water. In previous studies, a comparison of data from these 2 thermograph locations indicated that the 2 h hydraulic residence time in the streams resulted in slight heating or cooling of water in the streams depending upon ambient air temperatures (Culp and Podemski 1996). For example, over a 3 d period, the streams were cooler at night and warmer during the day as compared to the incoming river water. The maximum instantaneous difference between water temperature in the river and the streams was < 5 ^{0}C.

A repeated measured (RM) design is appropriate for analyzing physical and chemical water quality variables that are measured more than once. (Table 6), since measurements are made on the same replicates at each time (Paine, 1996). The RM design is useful for evaluating the effectiveness of the mesocosm approach in controlling all variables except those related to exposure to the effluent. Normally, in an impact or effects assessment, the analyses of the physical and chemical variables are considered supplementary or secondary to analyses of the biological and chemical endpoints (i.e., effects measures)

and are relegated to an appendix.

Table 11-6: ANOVA Table for Repeated Measure (RM) Design with Two Treatments, Four Replicate Streams per treatment, and t Measurements Times

Source		Df	F-test
Treatment	Tr	(Tr-1)	MS (Tr) /MS (Error 1)
Streams within treatments	Error 1	Tr(r-1)	MS (Error 1) /MS (Error 2)
Times	T	(t-1)	MS (T)/MS (Error 2)
Treatment x Times	Tr x T	(Tr-1)(t-1)	MS (Tr x T) /MS (Error 2)
Streams within treatments x Times	Error 2	Tr(r-1)(t-1)	

Where: df = degrees of freedom; MS Mean Square
Tr = number of treatments, r = number of replicate within treatments
t = number of repeated measurements

11.1.13 Exposure Endpoints

In addition to assessment of daily changes in physical water chemistry amongst all of the mesocosm streams, water samples should also be collected on a less frequent basis for analysis of general chemical parameters. It is recommended that samples be collected from the reference water head tank, from the full strength effluent, and from each mesocosm stream once at the end of the study. Samples should be analyzed for parameters as outlined in Environment Canada (supporting parameters) (2001). Often this information proves invaluable to confirm effluent dilutions in the treatment streams. For example, in freshwater systems exposed to pulping effluents, sodium is a relatively conservative ion that appears in high concentrations. By comparing Na^+ concentrations in the reference and exposure streams, the concentration of effluent in the streams can be verified.

Rhodamine dye has also been used on several occasions in mesocosm studies, following sampling of biota, to confirm effluent concentrations in the treatment streams. A concentrated solution of rhodamine is established in a storage tank and delivered to the treatment streams through the existing effluent lines. After a 24 hour period of turnover, water chemistry samples are collected in the rhodamine head tank and in each effluent treatment stream to calculate exposure concentrations.

The primary advantage of mesocosm systems is that they offer controlled exposure conditions. Thus chemical measurements to confirm exposure in biota are not normally required. However, these measures can provide valuable supporting information for site-specific assessments. For example, Dubé et al., (2001b) measured resin acids (i.e., dehydroabietic acid) in the bile of mummichog exposed to environmentally relevant concentrations of a mill effluent (3%) and compared these measures to control fish. Results showed significantly higher concentrations of DHAA in fish which confirmed that DHAA was a suitable tracer to confirm exposure to secondary treated pulp mill effluent where the mill furnish has a high proportion of softwood. These results also showed effluent-related differences between control and exposed fish.

Metal levels in indicator tissues have also been used as supporting information in mesocosm studies examining mine effluent effects on fish. Composite muscle tissue from juvenile Atlantic salmon showed significantly higher metal levels in fish exposed to mining effluent for 39 days compared to control fish (Dubé et al., unpublished). The metals measured in tissues were consistent with those detected in the effluent stream. The indicators and the tissues to measure metal exposure will vary with mine type and species being used. In general, gills, liver and kidney have the greatest potential for estimating exposure and bioavailability of metals. For small species (< 10 cm), whole body levels can be examined, or levels in the carcass after removing the digestive tract. Mercury is the only element that has been found to accumulate in muscle tissue, so if mercury is a contaminant of concern, dorsal muscle tissue should be analyzed. Blood and bone tissue may reflect exposure to lead, and might be considered if lead is the primary element of concern (Hodson et al., 1984). Bone concentrations are expected to be most indicative of long term metal exposure, while blood concentrations are indicative of short term exposure (AETE 1998).

11.1.14 Other Considerations

The guiding principles of the EEM program are that the program be scientifically defensible, cost-effective and provide flexibility for site-specific requirements. In order to keep the EEM programs focused and cost-effective, screening criteria were developed in the pulp and paper EEM program to ensure that only those monitoring techniques that were environmentally relevant, interpretable, and scientifically defensible were recommended in the EEM program (Table 7). Artificial stream systems satisfied all of the stated criteria with the following exceptions. For detecting the effects of effluents on fish, artificial stream studies do not examine multi-species interactions at a community of population level. However, it should also be recognized that the existing field surveys for fish also do not examine multi-species interactions at the community or population level.

Other limitations to application of artificial stream systems can include the lack of appropriate access and infrastructure requirements at some monitoring sites, and the restriction to medium term (i.e., weeks to months) monitoring such that long-term bioaccumulation processes cannot be examined.

Table 11-7: Evaluation criteria for EEM monitoring alternatives (after Dubé et al., 2001a).

Evaluation Criteria	Description
1. Environmentally Relevant	• Effects on relevant organisms are measured in the receiving environment.
2. Interpretable	• Measured endpoints are relevant to the detection/quantification of effects by pulp and paper effluents on fish, fish habitat or the use of fisheries resources. • Amenable to consistent scientific evaluation and interpretation. • An approach for determining "an effect" (statistically significant different between reference and exposed) has been demonstrated.
3. Scientifically Defensible and Manageable	• Acceptable test protocol exists, including QA/QC procedures. • Commercially available. • A validation study (e.g., pilot studies) has been completed.
4. Cost-effective	• The new technique provides equal or better quality data in comparison with currently accepted techniques. • Cost of the technique is similar to currently accepted techniques.

Cost-effectiveness of mesocosm application is critical for EEM application. Mesocosms can be an effective tool in a regulatory setting providing the system employed is cost-effective, and the design suited to the objectives of the program and the hypotheses to be tested. The costs for many components of a mesocosm study are the same as those for a field survey because the endpoints measured are the same. Laboratory analysis costs for example are the same. The main difference in cost between field surveys and mesocosm studies is in labor and use of the mesocosm facility itself. Labor costs and system transport costs depend on where the study is being conducted (i.e. site-specific) and the length of the exposure period. Transportation of the facility to the site, hook up of power and water, and colonization with biota are typically conducted within 4-5 days with a 4-person field crew and a mill electrician. Costs for system use and maintenance are supplier-specific. If suppliers have their personnel on-site for the duration of a 60 d exposure period then costs would be higher than a field survey because of supplier rates. Our approach has been to have crews for set-up and take-down, with day to day maintenance covered by mill personnel, or summer students, as these tasks take 1/2 day at most and do not require high levels of scientific expertise. As an estimate, mesocosm studies should fall within the range of field surveys + approximately 25%.

It must be recognized that mesocosms are alternatives to field surveys. If field surveys cannot be conducted then an alternative should be used to provide the required information. All efforts should be made to use alternatives in a cost-effective manner, as this is the most appropriate mechanism to attain quality data that is interpretable for confounded sites. If mesocosms are considered as an alternative for a particular mill, a further consideration to ensure cost effectiveness would be to incorporate more than one phase or more than one effluent (e.g., if joint studies between mills are possible) into the study design.

11.1.15 References

AETE (Aquatic Effects Technology Evaluation Program). 1998. Summary and cost-effectiveness of aquatic effects monitoring technologies applied in the 1997 AETE field evaluation program. Canada Centre for Mineral and Energy Technology, Ottawa, ON, Canada.

Anderson, A.M. 1989. An assessment of the effects of the combined pulp mill and municipal effluents at Hinton on the water quality and zoobenthos of the Athabasca River. Report, Environmental Quality Monitoring Branch, Environmental Assessment Division, Alberta Environment, Edmonton, AB, Canada.

Anderson, A.M. 1991. An overview of long-term zoobenthic monitoring in Alberta rivers (1983-1987). Report, Environmental Assessment Division, Alberta Environment, Edmonton, AB, Canada.

Barber, W.E., and N.R. Kevern. 1974. Seasonal variation of sieving efficiency in a lotic habitat. Freshwat. Biol. 4:293-300.

Buikema, A., and J.R. Voshell. 1993. Toxicity studies using freshwater benthic macroinvertebrates. In: D.M. Rosenburg and V.H. Resh (eds.), Biomonitoring and benthic invertebrates. Chapman and Hall, NY, USA, pp.340-395.

Burt, A., Kramer, D.L., Nakasuru, K. and C. Spry. 1988. The tempo of reproduction in *Hyphessobrycon pulchripinnis* (Characidae), with a discussion on the biology of "multiple spawning". Environ. Biol. Fish. 22: 15-27.

Carlisle, D.M. 2000. Bioenergetic food webs as a means of linking toxicological effects across scales of ecological organization. J. Aquat. Ecosystem Stress and Recovery 7:155-165.

Coad, B.W. 1995. Encyclopedia of Canadian Fishes. Canada Museum of Nature and Canadian Sportfishing Productions Inc. 928p.

Cooper, S. D., and L. Barmuta. 1993. Field experiments in biomonitoring. In: Rosenberg, D.M., and V.H. Resh (eds.), Biomonitoring and benthic invertebrates. Chapman and Hall, New York, NY, USA, pp.399-441.

Culp, J.M., and K.J. Cash. 1995. Potential impacts of effluent on the Thompson and upper Fraser Rivers. DOE FRAP 1995-10, Fraser River Action Plan, Environmental Conservation Branch, Environment Canada, Vancouver, BC, Canada.

Culp, J.M., and C.L. Podemski. 1996. Design and application of a novel stream microcosm system for assessing effluent impacts to large rivers. In: Servos, M.R., Munkittrick, K.R., Carey, J.H., and G.J. Van Der Kraak (eds.), Environmental fate and effects of pulp and paper mill effluents. St. Lucie Press, Delray Beach, FL, USA, pp.549-555.

Culp, J.M., Podemski, C.L., Cash, K.J., and R.B. Lowell. 1996. Utility of field-based artificial streams for assessing effluent effects on riverine ecosystems. J. Aquat. Ecosystem Health 5:117-124.

Culp, J.M., and R.B. Lowell. 1998. Pulp mill effluent impacts on benthic communities and selected fish species in the Fraser River Basin. In: Gray, C., and T. Tuominen (eds.), Health of the Fraser River Aquatic Ecosystem, Vol. II. DOE FRAP 1998-11, Environment Canada, Ottawa, ON, Canada, pp.13-34.

Culp, J.M., Lowell, R.B., and K.J. Cash. 2000a. Integrating mesocosm experiments with field and laboratory studies to generate weight-of-evidence risk assessments for large rivers. Environ. Toxicol. Chem. 19:1167-1173.

Culp, J.M., Podemski, C.L., Cash, K.J., and R.B. Lowell. 2000b. A research strategy for using stream microcosms in ecotoxicology: integrating experiments at different levels of biological organization with field data. J. Aquat. Ecosystem Stress and Recovery 7:167-176.

Culp, J.M., Podemski, C.L., and K.J. Cash. 2001. Interactive effects of nutrients and contaminates from pulp mill effluents on riverine benthos. J. Aquat. Ecosystem Stress and Recovery, in press.

Drake, J.A., Huxel, G.R., and C.L. Hewitt. 1996. Microcosms as models for generating and testing community theory. Ecology 77:670-677.

Dubé, M.G. and J.M. Culp. 1996. Growth responses of periphyton and chironomids exposed to biologically treated bleached kraft pulp mill effluent. Environ. Toxicol. Chem. 15:2019-2027.

Dubé, M.G. 2000. Sources of contaminants in the bleached Kraft pulping process and their effect on the mummichog (*Fundulus heteroclitus*). Ph.D. Thesis, University

of New Brunswick, Saint John, NB, Canada, 211p.

Dubé, M.G. and D.L. MacLatchy. 2000. Endocrine responses of an estuarine fish *Fundulus heteroclitus* to final effluent from a bleached kraft pulp mill before and after installation of reverse osmosis treatment of a waste stream. Environ. Toxicol. Chem 19:2788-2796.

Dubé, M.G. and D.L. MacLatchy. 2001. Identification and treatment of a waste stream at a bleached kraft pulp mill that depresses a sex steroid in the mummichog (*Fundulus heteroclitus*). Environ. Toxicol. Chem 20:985-995.

Dubé, M.G., Culp, J.M., Cash, K.J., Glozier, N.E., MacLatchy, D.L., Podemski, C.L., and R.B. Lowell. 2001a. Artificial streams for environmental effects monitoring (EEM): development and application in Canada over the past decade. Wat Quality Res J Can, in press.

Dubé, M.G., Culp J.C., MacLatchy, D.L., Gillis, G., Parker, R., Courtenay, S., and C. Gilman. 2001b. Utility of field-based artificial streams for assessing effects of pulp mill effluents on fish. J. Aquat. Ecosystem Stress Recovery, in press.

Environment Canada. 1992. Pulp and Paper Effluent Regulations, Canada Gazette, Part II, Volume 126, Number 11:1967-1997.

Environment Canada. 1995a. Further guidance for the invertebrate community survey for aquatic environmental effects monitoring related to Federal Fisheries Act Requirements. EEM 2, Ottawa. ON, Canada.

Environment Canada. 1995b. Further guidance for the adult fish survey for aquatic environmental effects monitoring related to Federal Fisheries Act Requirements. EEM 1, Ottawa, ON, Canada.

Environment Canada. 1997. Aquatic Environmental Effects Monitoring Requirements (revised EPS 1/RM/18). EEM/1997/1, Environment Canada, National EEM Office, Science Policy and Environmental Quality Branch Ottawa, ON, Canada.

Environment Canada. 1998. Pulp and Paper Technical Guidance for Aquatic Environmental Effects Monitoring. EEM/1998/1. Environment Canada, National EEM Office, Science Policy and Environmental Quality Branch, Ottawa, ON, Canada.

Environment Canada. 2001. Metal Mining Technical Guidance Document for Aquatic Environmental Effects Monitoring. EEM/2001/in press. Environment Canada, National EEM Office, Science Policy and Environmental Quality Branch, Ottawa, ON, Canada.

Forbes, V.E., and T.L. Forbes. 1994. Ecotoxicology in theory and practice. Chapman and Hall, New York, USA.

Fritz, E.S., Meredith, W.H., and V.A. Lotrich. 1975. Fall and winter movements and activity level of the mummichog *Fundulus heteroclitus*, in a tidal creek. Chesapeake Sci. 16: 211-215.

Gibbons, W.N., Munkittrick, K.R., and W.D. Taylor. 1998a Suitability of small fish species for monitoring the effects of pulp mill effluent on fish. 1. Response of spoonhead sculpin exposed to bleached kraft mill effluent. Environ. Toxicol. Chem., in press.

Gibbons, W.N., Munkittrick, K.R., and W.D. Taylor. 1998b. Suitability of small fish species for monitoring the effects of pulp mill effluent on fish. 2. Comparison of trout perch responses to white sucker exposed to bleached kraft mill effluent. Environ. Toxicol. Chem., in press.

Green, R.H. 1993. Application of repeated measures designs in environmental impact and monitoring studies. Austral. J. Ecology 18:81-98.

Heins, D.C. and F.G. Rabito Jr. 1986. Spawning performance in North American minnows: direct evidence of the occurrence of multiple clutches in the genus *Notropis*. J. Fish. Biol. 28: 343-357.

Hodson, P.V., B.R. Blunt, and D.M. Whittle. 1984. Monitoring lead exposure of fish. Pg. 87-98. *In:* Contaminant Effects of Fisheries. Cairns, V.W., P.V. Hodson, and J.O. Nriagu. John Wiley and Sons. New York, New York.

Jenkins, R.E. and N.M. Burkhead 1993. Freshwater fishes of Virginia. American Fisheries Society, Bethesda, ML, USA, 1079p.

Jonasson, P.M. 1955. The efficiency of sieving techniques for sampling freshwater bottom fauna. Oikos 6:183-207.

Kneib R.T., and A.E. Stiven. 1978. Growth, reproduction, and feeding of *Fundulus heteroclitus* (L.) on a North Carolina salt marsh. J. Exp. Mar. Biol. Ecol. 31:121-140.

Leblanc, J. et C. M. Couillard. 1995. Description de la période de reproduction d'un poisson sentinelle: le choquemort (Fundulus heteroclitus) de l'estuaire de la Miramichi Rapp. tech. can. sci. halieut. aquat. 2057: vii + 32p.

Leblanc, J., Coulliard, C.M., and J.F. Brethes. 1998. Modifications of the reproductive period in mummichog (*Fundulus heteroclitus*) living downstream from a bleached

Kraft pulp mill in the Miramichi Estuary, New Brunswick, Canada. C.J.F.A.S., in press.

Lowell R.B., Culp J.M., and F.J. Wrona. 1995. Stimulation of increased short-term growth and development of mayflies by pulp mill effluent. Environ. Toxicol. Chem. 14:1529-1541.

Lowell R.B., Culp J.M., Wrona F.J., and M.L. Bothwell. 1996. Effects of pulp mill effluent on benthic freshwater invertebrates: food availability and stimulation of increased growth and development. In: Servos, M.R., Munkittrick, K.R., Carey, J.H., and G.J. Van Der Kraak (eds.), Environmental fate and effects of pulp and paper mill effluents. St. Lucie Press, Delay Beach, FL, USA, pp.525-532.

Lowell R.B., Culp J.M., and M.G. Dubé. 2000. A weight-of-evidence approach for northern river risk assessment: integrating the effects of multiple stressors. Environ. Toxicol. Chem. 19:1182-1190.

McIntire, C.D. 1993. Historical and other perspectives of laboratory stream research. In: G.A. Lamberti and A.D. Steinman (eds.), Research in artificial streams: applications, uses and abuses. J. North Amer. Benth. Soc. 12:313-323.

Nelson, J.S. and M.J. Paetz. 1992. The fishes of Alberta. The University of Calgary Press, Calgary, AB, Canada.

Paine, M.D. 1990. Life history tactics of darters (*Percidae: Etheostomatiini*) and their relationship with body size, reproductive behaviour, latitude and rarity. J. Fish Biol. 37: 473-488.

Paine, M.D. 1996. Repeated measures designs. Environ. Toxicol. Chem. 15:1439-1441.

Peters, R.H. 1983. The ecological implication of body size. Cambridge University Press, New York, USA, 329p.

Podemski C.L., and J.M. Culp. 1996. Nutrient and contaminant effects of bleached kraft mill effluent on benthic algae and insects of the Athabasca River. In: Servos, M.R., Munkittrick, K.R., Carey, J.H., and G.J. Van Der Kraak (eds.), Environmental fate and effects of pulp and paper mill effluents. St. Lucie Press, Delay Beach, FL, USA, pp.571-580.

Podemski C.L. 1999. Ecological effects of a bleached kraft pulp mill effluent on benthic biota of the Athabasca River. Ph.D. Thesis, University of Saskatchewan, Saskatoon, SK, Canada, 292 p.

Rees, H.L. 1984. A note on mesh selection and sampling efficiency in benthic studies. Mar. Poll. Bull. 15:225-229.

Resh V.H., and E.P. McElravy. 1993. Contemporary quantitative approaches to biomonitoring using benthic macroinvertebrates. In: Rosenberg, D.M. and V.H. Resh (eds.), Freshwater biomonitoring and benthic invertebrates. Chapman & Hall, NY, USA, pp.159-194.

Roberts, W.E. 1988. The sculpins of Alberta. Alberta Naturalist 18:121-127.

Scott, W.B. 1967. Freshwater fishes of eastern Canada. University of Toronto Press, Toronto, ON, Canada. 137p.

Scott, W.B. and E.J. Crossman, 1973. Freshwater fishes of Canada. Fish. Res. Board. Can. Bull. 184. Ottawa, ON, Canada. 966p.

Scrimgeour, G.J., Culp, J.M., and N.E. Glozier. 1993. An improved technique for sampling lotic invertebrates. Hydrobiologia 254:65-71

Shuter, B.J. 1990. Population-level indicators of stress. Amer. Fish. Soc. Symposium 8:145-166.

Slack, K.V., Averett, R.C., Geeson, P.E., and R.G. Lipscomb. 1973. Methods for collection and analysis of aquatic biological and microbiological samples. In: Techniques of water-resources investigations of the United States Geological Survey, Chapter 4A, Book 5, U.S. Department of the Interior, Geological Survey, Washington, D.C., USA, pp. 1-165.

Snedecor, G.W. and W.G. Cochran. 1980. Statistical methods, 7th ed. Iowa State University Press, Ames, Iowa, USA, 507p.

Stewart-Oaten, A., Bence, J.R., and C.W. Osenburg. 1992. Assessing effects of unreplicated perturbations: no simple solutions. Ecology 67:929-940.

Suess, M.J. 1982. Examination of water for pollution control. A reference handbook. Vol. 3. Biological, bacteriological and virological examination. Permagon Press, Oxford, England.

Taylor B.R., and R.C. Bailey. 1997. Technical evaluation on methods for benthic invertebrate data analysis and interpretation, final report. AETE project 2.1.3. Natural Resources Canada, CANMET, Ottawa ON, Canada, 26p.

Weber, C.I. 1973. Biological field and laboratory methods for measuring the quality of surface waters and effluents. EPA-670/4-73-001. U.S. Environmental Protection Agency, Cincinnati, OH, USA.

Wiederholm, T. 1980. Use of benthos in lake monitoring. J. Wat. Poll. Cont. Fed. 52:537-547.

Wrona, F.J., J.M. Culp and R.W. Davies. 1982. Macroinvertebrate subsampling: A simplified apparatus and approach. Can. J. Fish. Aquat. Sci. 39:1051-1054.

11.2 Use of Caged Bivalves as Alternative Monitoring Method

11.2.1 Introduction

A description of methods for caged bivalve studies is provided in this section, which includes detailed guidance on:

- background and general approach of caged bivalve studies;
- species selection;
- study design;
- variables to be measured;
- methods for implementing the study;
- data analyses;
- tissue concentrations; and
- reporting requirements.

11.2.2 Background

In October 2000, the National Environmental Effects Monitoring (EEM) Science Committee recommended the use of caged bivalve studies as a scientifically defensible alternative approach to a wild fish survey if it is not practical or technically feasible to conduct the fish survey (FS), according to the decision tree in Annex 1 of the Pulp and Paper EEM Technical Guidance Document (Environment Canada 1998: Figure 5.1). The Metal Mining Fish Subgroup also recommended the use of caged bivalves as an alternative method during the Metal Mining EEM multi-stakeholder consultation process.

During the first cycle (1993 to 1996) of the pulp and paper EEM program, the FS performed reasonably well in the freshwater environment, but a number of problems were encountered for marine and estuarine environments. In February 1997, a workshop was convened to discuss alternatives to the FS component of the EEM program (Courtenay et al. 1998). One of the suggested alternatives was caged bivalve studies. Subsequently, two pilot studies using caged bivalves were conducted for marine environments: in 1997 at Port Alice on Vancouver Island (Applied Biomonitoring 2000) and in 1998 at Pictou Harbour in Nova Scotia (Andrews and Parker 1999). Based on these two pilot studies and several other related studies using caged bivalves, the Science Committee and Metal Mining Fish Subgroup determined that an EEM study using caged bivalves can produce good quality data on effluent effects on fish that fit into the regulatory context.

This technical guidance describes procedures for conducting controlled experiments using caged bivalves suspended in the water column, to test for effects associated with

industrial discharges, comparing measurements from an exposure area to reference areas. This approach is an alternative for the FS and may be considered for mills where the FS has been unsuccessful or impractical in past cycles of EEM, or where there may be study design issues, such as confounding influences, or safety concerns.

In November 2000, American Society for Testing and Materials (ASTM) approved a method for conducting environmental studies using caged bivalves (Salazar and Salazar 2000). The ASTM method serves as the basis for this technical guidance on conducting caged bivalve studies. A number of other studies using caged and wild bivalves were also considered in developing this specific guidance for applying this approach within the framework of the EEM program. The format of this guidance follows that used for the FS.

Bivalves are useful biomonitors for several reasons:

- bivalves are relatively non-mobile, such that exposure to contaminants is assured and representative of the exposure area;
- bivalves are abundant in many marine, estuarine and freshwater environments, and are relatively easy to handle and sample year-round;
- the biology of many shellfish species is well known and considerable research has been conducted on effects to shellfish from exposure to various anthropogenic and natural environmental stressors;
- several bivalve species have been shown to readily accumulate many chemicals from a variety of pathways (water, sediment, food) and show sublethal effects associated with exposure;
- bivalve growth is relatively easy to measure and has been shown to be as sensitive or more sensitive than mortality in other standard test species such as *Daphnia*, fathead minnow and rainbow trout (see Salazar and Salazar 2000); and
- bivalves are an important fisheries resource and both the Atlantic and Pacific regions support commercially valuable shellfish aquaculture industries, as well as commercial and recreational shellfish harvests.

Bivalves, such as oysters and mussels, have been used in Mussel Watch programs since the mid-1970s to monitor trends in chemical contamination and assess the effects of human activities on coastal and estuarine areas. Mussel Watch programs began in the U.S. (Goldberg et al. 1978) and have since become international in scope (Jernelov 1996).

11.2.3 Approach

The caged bivalve approach provides a number of advantages to the investigator in conducting a monitoring program. These include experimental control and realism, use of

organisms naturally found in the study environment, and known exposure period. By using caged bivalves, rather than resident populations, the variability in biological measurements can be reduced by using individuals of similar size and environmental history, thereby increasing the discriminating power of the test (Salazar and Salazar 1995). A considerable number of caged bivalve studies have been conducted in Canada and the U.S., as well as other countries (see Salazar, in Stewart and Malley 1997).

The target endpoints for caged bivalve studies in EEM are primarily survival, growth and condition. Reproduction is considered a key endpoint for fish population monitoring, but currently there are no standardized methods to monitor this in caged bivalves. Once these methods are established, this document will be updated to include guidance on this. A tissue analysis (dioxins and furans) may be required, and caged bivalves can be used to meet this requirement. Other chemicals or metals may be used to assess bioaccumulation to aid in interpreting results or used during investigation of cause.

There are some limitations to using caged bivalve studies as surrogates for fish surveys. These include (from Salazar, in Stewart and Malley 1997):

- possible effects of transplanting causing stress to bivalves;
- loss of cages from acts of nature, inadvertent capture by moving vessels, vandalism; and
- costs associated with collection, sorting and deployment.

That being stated, the caged bivalve approach is considered to be a potential alternative in some cases where a wild fish survey cannot be successfully conducted.

11.2.4 Species Selection

Many bivalve species have been used for assessing chemical bioavailability or effects in marine, estuarine and freshwater environments.

Species selection for caged bivalve studies should be made carefully and consider at least the following questions:

- Are conditions at the exposure and reference areas similar to the natural habitat of the species in terms of tolerance limits for natural factors such as temperature, salinity, dissolved oxygen, and pH?
- Is the species naturally present in the area under evaluation?
- Is there documentation that the species can accumulate and/or be sensitive to the contaminants of concern?
- Is the life history of the species well known in terms of spawning cycle and life-stage requirements?
- Does the species have threatened or endangered status?

- Is an abundant supply of the species readily available?
- Is the species easy to handle in the field?

Ideally, species should be used that have wide geographic distributions so that test results can be compared across studies.

11.2.4.1 Most Commonly Used Taxa

The species most commonly used in field bioassays in Canadian waters and considered relevant for use in the EEM programs are described below and listed in Table 8. The temperature and salinity tolerance limits for each species are provided, along with information on age at maturity, spawning periods and general distribution within Canada. Other species, such as marine clams *(Mya arenaria, Macoma balthica)* and scallops, may also be suitable for some applications. However, the environmental requirements and sensitivity of alternative bivalve test species should be established before they are used in an EEM program.

Table 11-8: Suggested Taxa for Use in Caged Bivalve Studies for EEMs

Species and Reference	Temperature Range (°C)	Salinity Range (parts per thousand)	Reproductive Information	General Distribution in Canada
Marine and Estuarine Bivalves				
Mytilus edulis (blue mussel) (Freeman et al. 1994; Grout and Levings 2000; Mucklow 1996; Stewart 1994; Salazar and Salazar 2000; Newell 1989)	-1.5 to 25	5 to 33	Matures at 2 years of age. Spawning varies among populations; some low-level throughout year; first in early summer, second in the fall	Atlantic coast
Mytilus trossulus (Pacific blue mussel) (Freeman et al. 1994; Salazar and Salazar 2000; Skidmore and Chew 1985)	0 to 29	4 to 33	Matures at 2 years of age. Spawning from July to September	Atlantic and Pacific coasts
Crassostrea gigas (Pacific oyster) (Virginia Tech 2000; Waldock et al.1996)	4 to 24	25 to 35	Spawning July to August.	Pacific coast
Freshwater Bivalves				
Elliptio complanata (Eastern elliptio) (Beckvar et al. 2000; Day et al. 1990; Hinch and Green 1989; McMahon 1991; Metcalfe-Smith et al. 1996)	0 to 30	0 to 3	Age at maturity 6-12 years. Spawning occurs mostly June to July; some May to September.	Eastern Canada
Pyganodon (Anodonta) grandis (common floater clam) (Clarke 1973; Couillard et al. 1995a; 1995b; Malley et al. 1996)	0 to 30	0 to 3	Spawns mostly April to May, some to August	Interior and Eastern Canada
Anodonta kennerlyi (western floater clam) (BC Research Inc. 1998; Clarke, 1981; Stewart and Malley 1997; Williams et al. 1993)	0 to 30	freshwater	Spawning begins in early August.	Alberta and British Columbia
Sphaeriid clams (e.g., *Musculium securis, Sphaerium rhomboideum, S. striatinum*) (Hornbach et al. 1982; Mackie 1978a; 1978b; Mackie and Flippance 1983; Mackie et al. 1974; Stephenson and Mackie 1981)	10 to 25* (* optimal growth range)	freshwater	Life cycle generally #1 year; life histories of many species are well documented; reproductive effort can be quantified.	Widely distributed in Canada

(Format modified from Salazar and Salazar 2000.)

11.2.4.2 Marine and Estuarine Bivalves

Mytilus sp. is an intertidal bivalve that has been successfully used in caged studies for marine and estuarine environments since the late 1970s. *Mytilus* has been used extensively in the ongoing International Mussel Watch Project to monitor trends of chemical contamination and assess the effects of human activities on coastal and estuarine areas in North America and around the world (O'Connor 1992; Jernelov 1996). The blue mussel, *M. edulis*, and *M. trossulus* (Pacific blue mussel) are found on the Atlantic Canada coast, whereas only *M. trossulus* is found on the Pacific Canada coast. These two species may be easily confused where they co-occur on the Atlantic coast (Freeman et al. 1994; Mucklow 1996). *M. edulis* was successfully used for the caged bivalve pilot studies for pulp and paper EEM on the Pacific coast (Applied Biomonitoring 2000) and the Atlantic coast (Andrews and Parker 1999) and has otherwise been widely used as a biomonitor (Stewart and Malley 1997; Salazar and Salazar 2000). *M. trossulus* has been successfully used in a number of caged bivalve studies in Pacific coastal waters (Stewart and Malley 1997).

Crassostea gigas is an oyster species that has been used in transplant studies in marine and estuarine studies (Waldock et al. 1996). The shells of *C. gigas* are usually more difficult to measure because of their irregular shape and protrusions.

11.2.4.3 Freshwater Bivalves

Unionid bivalves (such as *Elliptio complanata* and *Pyganodon grandis* (formerly *Anodonta grandis*)) are freshwater unionid mussels that have been used in a number of caged studies to examine water column and sediment exposures. *Anodonta kennerlyi* is a freshwater unionid mussel that was recently used for a caged mussel study on mercury uptake in a reservoir in BC (BC Research 1998).

A note of caution should be raised regarding the use of unionids for caged studies as part of EEM programs. Unionids have been used in a number of caged studies in which the endpoint of interest is bioaccumulation, whereas the focus for this technical guidance is survival, growth and condition. Obtaining sufficient numbers for use in a caged study may be very difficult and may jeopardize the status of natural unionid populations. Unfortunately, they are not commercially grown and are therefore not available in cultured form. Unionids are a long-lived group of organisms with slow recruitment. In the wild, they are often found as mature, large-sized, slow growing individuals and smaller, faster growing individuals are notoriously difficult to find.

As an alternative to unionids, sphaeriid clams might be considered suitable for assessing differences in the survival, growth, condition and reproductive rates of bivalves in freshwater receiving environments. Sphaeriid clams have short life cycles (generally one year), are abundant and widespread in distribution, and do not depend on fish hosts for a

part of their life cycle (as unionids do). Sphaeriids are also generally more tolerant than unionids of low calcium concentrations in the environment (Stewart and Malley 1997), which may be an important factor in many regions of Canada. Sphaeriid clams are ovoviviparous (Mackie 1978a) and hold neonate clams in brood pouches on their gills until they are ready to be released (Mackie et al. 1974). Sphaeriid clams have been successfully used in the past to assess the effect of pulp mill effluent and other chemical stressors on growth and natality (Mackie 1978b; Sandusky and Sparks 1979; Stephenson and Mackie 1981). Potentially useful species for EEM studies include various genera of *Sphaerium* (particularly *S. striatinum*) and *Musculium* (particularly *M. securis*). The various taxa of the genus *Pisidium* may be too small for practical handling and, in addition, are taxonomically difficult for the non-specialist. Additional research may be required to demonstrate the utility of sphaeriid bivalves for use in EEM.

11.2.4.4 Source

Bivalves may be obtained either from natural populations or from culturing facilities. Where possible, bivalves should be obtained from a commercial grower who can give certification or other assurance that the supplied culture is single-species and taxonomy is confirmed. Animals obtained from a commercial facility have an environmental history that is well known and assurances of being uncontaminated are greater than for animals collected in the wild. In any study, some form of certificate should be provided to assure the identity of the species used (i.e., a letter from the supplier or a taxonomist's letter of validation).

All individuals used in a caged bivalve study should be from the same population, because different populations of the same species may differ in their response to the exposure. If wild populations are the only possible source, then they must be collected from an uncontaminated area. For wild populations, infaunal bivalves can be separated from sediment by hand or by gentle sieving. Care should be taken not to over-sample an existing natural population. If the species of choice uses byssal threads for attachment to substrate, a knife or scissors should be used to gently remove byssal threads from the outside of the shell so that the internal byssal glands are not damaged. Remove all epiphytic growth from bivalve shells gently by hand or by gentle use of a soft brush or scraper.

Collection permits for field-collected bivalves may be required by some local, provincial and federal agencies. Transplant or aquaculture permits may be required in advance from local and provincial agencies. In addition, if cages present a potential obstruction to navigation, then a permit may be required from Canadian Coast Guard and a Notice to Mariners may be required. The permitting process should be considered early in the study planning process, as it may take several weeks to months to process the necessary permits.

11.2.4.5 Size and Age

All bivalves used in a caged study should be from the same age class and as uniform as possible in size. Age class is more difficult to determine in wild specimens than in cultured specimens because wild populations may consist of several age classes, with individuals growing at different rates. Sub-adult (immature) specimens may be selected because they typically have the greatest potential for growth. On the other hand, it may be desirable to use adults that have just completed spawning so that growth during the exposure period reflects growth in both somatic and reproductive tissues. It is important to choose specimens that will not become reproductive and spawn during the exposure period because spawning may confound the results. Investigators should attempt to note the reproductive status of a subsample of the test specimens at the beginning and end of the test and note whether spawning occurred.

11.2.4.6 Number of Organisms

Power analysis should be used to determine the minimum number of bivalves to detect a specified "effect size". It is recommended that studies be designed to detect a 20 percent change in growth. Data sets to aid in predicting the number of required organisms may be available through Environment Canada.

The number of animals required per cage for the growth endpoint will depend on study design (e.g., number of cages per station; see more discussion on this in Section 11.2.7.1), species used, age of animals, variability in response to the station, and growth conditions at the station. The number of bivalves required to fill the cages will depend on study design in terms of :

- number of areas (e.g., exposure area plus number of reference areas);
- number of stations per area;
- number of moorings per station;
- number of cages per mooring (if the study is designed to address varying depths in the water column); and
- bivalves per cage.

In addition, if tissue samples are required for tissue analysis, then consideration must be given to the number of animals required to obtain a sufficient sample size for chemical analysis.

When the minimum number of animals required to achieve the desired power has been estimated, the number of bivalves collected needs to be determined, and should include:

- total number required for test set-up (see set of bullets above);

- plus the number required for baseline measurements at the beginning of the test;
- plus the number required to account for dead or dying individuals during transport and handling (dead or dying bivalves are identified by smell or if their shells do not close very quickly in response to tapping or do not close at all); and
- plus 50 to 100 percent more to account for size variability and allow you to obtain the size fraction you are targeting.

11.2.4.7 Handling and Holding Conditions

Salazar and Salazar (2000) provide detailed guidance on handling and holding conditions for bivalves, and this is summarized below. Test organisms should be handled as little as possible and should be deployed as soon as possible after collection. When handling is necessary, it should be done carefully, gently and quickly so that the bivalves are not unnecessarily stressed. Bivalves should be kept in well-aerated, clean-flowing water as long as possible between collection, sorting and deployment. If transporting bivalves for extended periods, keep them moist and cool by placing them in a cooler with frozen gel packs or wet ice (wet ice at the bottom of the cooler). Use newspaper or cloth towels to keep bivalves separated from the gel packs or wet ice.

For use of caged bivalves in EEM, there may be concerns about bivalves coming into contact with chemicals used in gel packs, chlorine present in wet ice, or other chemicals that may be released from newspaper and towels. The best strategy to avoid the need for holding at low temperature is to minimize time between collection and deployment. Bivalve supply should be as close to the deployment sites as possible. If there are extended periods (over one day) between collection and deployment, bivalves should be held in clean (uncontaminated), well-aerated water kept at the cool end of the seasonal environmental range for the test species. Placing small containers with bivalves within a larger container in which temperature is controlled should maintain bivalves at the desired temperature.

11.2.5 Study Design

Determining the appropriate study design is critical if the results are going to be meaningful. Study design for caged bivalves includes a number of components:

- sampling design;
- area and station selection;
- replication of cage stations and cages per station;
- timing and duration;
- modifying or confounding factors;

- supplementary measures;
- cage design; and
- mooring systems.

11.2.6 Sampling Design

There are five main sampling designs that are recommended for marine and freshwater assessments:

- Control-Impact (C-I) design;
- Multiple Control-Impact (MC-I) design;
- Simple Gradient (SG) design;
- Radial Gradient (RG) design; and
- Multiple Gradient (MG) design.

The C-I and MC-I designs are used to determine the magnitude of difference between homogeneous exposed and unexposed areas, while the SG, RG and MG designs examine changes in an effect along an effluent gradient. The C-I and MC-I designs are used when there are few qualitative levels of exposure. It is suggested that multiple reference areas be used rather than increasing sample sizes in one reference station (Environment Canada 1998). The SG, RG and MG designs may be used when there are many quantitative levels of exposure (Paine 2000) and also may be useful in discriminating among effects from sources other than pulp mill effluent. Guidance on selecting the appropriate design is provided in Sections 5.1.3.3 and 6.1.5.4 of the pulp and paper Technical Guidance Document (Environment Canada 1998) and Section 5.6 of the metal mining Guidance Document (Environment Canada 2001).

For example, in estuaries with complex tidal regimes and mixing regimes, an MC-I approach may be the most appropriate. A SG approach may be applicable to a river receiving environment where flow is unidirectional.

C-I and MC-I designs require some level of replication within the control and impact areas. Replication at each station may not be required for gradient designs, although an appropriate number of stations is required in order to discriminate between spatial patterns related to effluent discharge, and other spatial patterns in the environment. Replication is discussed below in Section 11.2.7. Ultimately, it is the responsibility of study designers to develop site-specific study designs that are scientifically defensible, robust, and suitably sensitive.

11.2.6.1 Area and Station Selection

Section 3 of the pulp and paper Technical Guidance Document (Environment Canada 1998) and Chapter 3, Section 4.2.2 and Section 5.8 of the metal mining Guidance Document (Environment Canada 2001) provides guidance for selection of multiple reference locations for a variety of receiving environments and is applicable to caged mussel studies. Reference areas should be as similar as possible to study areas in terms of:

- hydrodynamic conditions;
- temperature;
- salinity;
- dissolved oxygen; and
- food availability and quality.

Multiple reference stations may help to identify natural differences and variability among uncontaminated areas.

The exposure area is defined by plume characteristics. Plume delineation, as described in Section 2.2 of the pulp and paper Technical Guidance (Environment Canada 1998) and Section 3.3 in the metal mining Guidance Document (Environment Canada 2001), should provide sufficient information to model average effluent concentrations with distance from source and to identify the degree of vertical mixing in the water column. This information will assist in the selection of stations and depth of deployment in the water column. More guidance is provided below in Section 11.2.7.1.

Boating and recreational activities in exposure and reference areas should be considered during station selection to minimize the incidence of cage removal or damage due to boating activities and to minimize incidents of vandalism. Boat basins and areas with significant marine traffic may also result in exposure of bivalves to petroleum products, anti-foulants and other chemicals associated with boating activities.

11.2.7 Replication

11.2.7.1 Number of Cage Stations

For the exposure area, stations and depths for cage deployment may be chosen to represent a gradient of exposure to the effluent plume. The deployment stations of bivalve cages should be given careful consideration for a number of potentially interfering factors, such as.

- in estuarine and marine situations, effluent may be positively buoyant, resulting in a thin layer of "freshwater" effluent floating on denser saline water;

- in tidal situations, the mixing behaviour of effluent may be quite complex, resulting in low confidence about the average exposure concentration that bivalves may be exposed to. Tidal situations may require consideration of exposure area stations that are both "upstream" and "downstream" of the outfall; and
- in river discharge situations, effluent plumes may be quite long and narrow, meaning that there is little or no opportunity to meaningfully replicate cage stations within (i.e., across the long axis of) the plume.

For these and other reasons articulated below, it might be more meaningful to simply evaluate "distance from outfall" rather than effluent concentrations when designing caged bivalve studies. This approach would remain consistent with one of the objectives of EEM, which is to evaluate the magnitude and geographic extent of effects that may be related to the discharge of effluent.

11.2.7.2 Replication and Pseudoreplication

The number of replicate stations must be determined to address the sampling design and EEM objectives. The number of replicate stations and subsamples within replicate stations will be determined by power analysis. The allocation and distribution of replicate stations is dependent upon the sampling design. Guidance on the use of power analysis is provided in Section 6.1.5.5 of the pulp and paper Technical Guidance Document (Environment Canada 1998) and Sections 4.4.1.1 and 5.11 of the metal mining Guidance Document (Environment Canada 2001).

Any discussion of replication in a field study must consider the question of "pseudoreplication", defined by Hurlbert (1984) as "the use of inferential statistics to test for treatment effects with data from experiments where either treatments are not replicated (though samples may be) or replicates are not statistically independent". In ANOVA terminology, it is "the testing for treatment effects with an error term inappropriate to the hypothesis being tested" (see Box 1 for further discussion). See also Section 3 of the Technical Guidance Document (Environment Canada 1998) for further discussion on Hurlbert (1984) and the issue of pseudoreplication.

The issue of pseudoreplication is one that will continue to challenge field experiments for the foreseeable future. For EEM studies, where a single effluent stream is discharged into the receiving environment, it could be argued that no true replication is possible, and that any set of samples must be correlated with each other in ways that undermine inferential statistics. The onus will remain on study designers to develop approaches to EEM studies that minimize the risk of drawing incorrect conclusions as a result of the potentially "misleading" side-effects of pseudoreplication (see Table 1 in Hurlbert 1984).

It would appear that the best approach would be to evaluate (using analysis of variance (ANOVA) or other statistical techniques) whether there are significant differences between stations that happen to be strategically located around an effluent outfall. Any

interesting correlation that may exist between significant differences and distance from the outfall (or the expected effluent concentration) may also be noted. It may be necessary to conduct further studies to attribute measured differences to the effluent.

This limitation is well recognized within the overall EEM context, and applies with equal force to FS and benthic invertebrate community survey.

Box 1 Discussion of Pseudoreplication in the Context of Hurbert (1984)

The discussion of pseudoreplication must refer back to one of the objectives of the EEM studies, which is to measure the *magnitude and geographicl extent* of any effects of effluent on specific aspects of the receiving environment and/or its biota. There is an important point implicit in this objective, and that is that effects we wish to detect are unique to the mill being investigated. Every combination of mill and receiving environment is unique, therefore care must be taken when generalizing to other receiving environments with respect to environmental effects associated with a single effluent. Also implicit in the objective of the EEM studies is that we wish to evaluate whether there is a statistically significant difference in the environment or biota caused by mill effluent at various locations near the effluent outfall.

Returning to Hurlbert (1984), two classes of experiments are defined: *mensurative* (involving only the making of measurements at one or more points in space or time), and *manipulative* (involving two or more treatments, having as a goal the making of one or more comparisons, and organized so that different experimental units receive different treatments, and that the assignment of treatments to experimental units is or can be randomized). In the EEM context, we choose to regard locations as experimental units, rather than cages containing one or more bivalves. In making this decision, we regard the EEM studies to be mensurative in character, with the objective of assessing whether a biological property (such as survival or growth of bivalves) is different at one location than at another.

Hurlbert (1984) provides as illustrations several hypothetical examples of experiments. We refer specifically to Hurlberts's *Example 3*, where the decomposition of leaves contained in a number of mesh bags is being compared at two different locations (the 1 m and the 10 m isobath of a water body). Hurlbert notes (italics his):

"....out of laziness, we place all eight bags at a single spot on each isobath. It will still be legitimate to apply a significance test to the resultant data. However, and the point is the central one of this essay, if a significant difference is detected, this constitutes evidence only for a difference between the two (point) *locations*; one 'happens to be' a spot on the 1-m isobath, and the second 'happens to be' a spot on the 10-m isobath. Such a significant difference cannot legitimately be interpreted as demonstrating a difference between the two *isobaths*, i.e., as evidence of a 'treatment effect'. For all we know, such an observed significant difference is no greater than we would have found if the two sets of eight bags had been placed at two locations on the *same* isobath.

If we insist on interpreting a significant difference in Example 3 as a "treatment effect", then we are committing what I term *pseudoreplication*."

For more discussion about pseudoreplication, please consult Hurlbert (1984).

11.2.7.3 Applying Replication To Study Design

Several measurement endpoints are suggested for caged bivalve studies (survival; growth -change in length or wet weight; soft tissue fresh weight; condition). In order to assess these endpoints on a set of bivalves exposed at different locations in the field, many different configurations of cages and bivalves are possible. One possible approach is outlined below and illustrated in Box 2, although it is emphasized that site-specific conditions may make a wide variety of different options viable.

Assume sets of 100 bivalves are deployed at a reference area, and in the exposure area at distances of 50, 200 and 400 m away from an effluent outfall. A hypothetical data set was generated using random numbers, representing this scenario; the data can be considered to represent the soft tissue fresh weight of bivalves at the end of the exposure period (see Box 2). To generate this data set, four sets of 100 random numbers were generated, with the sets having expected mean values of 13 (reference), 10 (50 m), 11 (200 m) and 12 (400 m), and standard deviations of 2.0. Each set of 100 numbers was subsequently split into five groups of 20, representing bivalves deployed on individual moorings. Note that this approach assumes that moorings are replicates, and that the individual bivalves deployed on each mooring are sub-samples (not true replicates).

Three potential study designs are evaluated using these data. In the first design, all 100 bivalves at each Distance are considered to be replicates (although this is technically incorrect), and the resulting ANOVA has 3 degrees of freedom (DF) for Distance, and 396 DF for experimental error. The difficulty with this design is that only a single estimate of survival is available at each Distance, although tests of significance for the effect of area on growth are valid (ANOVA 1).

In the second design (ANOVA 2), bivalves are deployed in cages of 20 individuals, and each bivalve within each cage is treated as an individual sub-sample. In this analysis, five estimates of survival are available for each Distance, and the ANOVA (which separates out the Subsampling Error and considers the mean square error for "Cages Within Distance" to be the appropriate error term for testing effects of Distance) is more powerful than the analysis in ANOVA 1, despite being based on the same data.

In the third design (ANOVA 3), bivalves are again deployed in cages of 20 individuals, but for each cage, only the mean growth rate within cage is estimated. In this analysis, the test for effects of Distance has exactly the same power as the analysis using subsampling.

Box 2: Bivalve tissue dry weight as a function of distance from mill outfall (synthetic data)

50 m					200 m				
14.53	11.77	10.38	8.70	11.06	8.65	10.74	9.07	8.34	12.13
6.44	10.66	10.71	6.90	12.25	14.69	8.34	9.78	11.57	10.47
9.43	9.40	11.47	9.54	11.33	10.46	11.57	10.64	15.26	9.33
10.22	10.07	11.77	7.44	7.86	11.76	8.92	10.87	12.11	11.12
9.13	5.99	9.11	7.86	8.62	10.43	13.75	7.98	12.03	13.90
9.51	10.58	9.90	10.50	10.50	12.51	11.62	13.53	12.95	13.42
11.32	7.63	10.70	13.17	10.88	12.62	12.38	11.12	7.03	8.70
9.21	10.38	14.57	6.26	10.11	11.97	11.56	12.73	12.82	10.75
14.79	6.91	7.21	10.31	7.92	10.15	6.87	13.77	12.57	11.52
8.68	8.23	10.85	7.19	9.61	13.85	10.65	12.96	7.28	12.20
9.49	13.16	9.62	10.68	9.20	14.27	13.75	7.39	9.57	7.22
8.49	8.18	6.70	7.30	8.81	9.99	10.24	12.02	9.87	8.65
11.13	7.66	14.63	7.76	7.92	10.21	11.11	9.42	9.36	14.28
12.07	10.39	7.84	12.03	12.61	11.35	8.36	12.80	9.03	10.95
13.27	9.32	12.03	10.66	10.71	7.81	7.08	13.55	11.46	15.76
11.75	10.81	5.43	7.91	7.69	9.93	10.94	12.44	11.71	12.31
9.72	12.35	8.52	10.19	10.75	14.50	12.05	8.77	10.02	13.71
9.61	7.63	13.00	8.52	13.98	9.56	9.35	8.75	10.85	7.08
10.12	10.73	10.41	10.27	10.40	9.02	12.39	10.66	14.98	9.76
5.65	9.07	11.27	9.89	9.33	9.94	10.49	12.71	10.61	12.78
10.228	**9.546**	**10.306**	**9.154**	**10.077**	**11.184**	**10.608**	**11.048**	**10.971**	**11.302**

400 m					Reference				
11.02	11.01	8.69	10.90	9.93	17.18	15.75	7.76	15.34	11.59
13.22	10.99	9.91	12.00	13.77	13.22	13.72	12.84	13.87	10.75
12.30	13.98	13.31	11.14	12.09	11.77	12.47	12.29	11.99	10.54
9.10	10.32	9.48	11.36	12.38	14.58	14.02	9.73	11.85	14.51
12.12	12.30	12.34	14.42	11.12	13.11	16.24	12.80	12.50	15.52
11.90	12.00	9.28	8.51	11.80	13.60	12.65	11.64	12.37	11.31
14.33	12.13	12.15	9.40	11.57	15.27	11.78	13.77	10.25	14.00
10.32	9.19	11.82	14.16	10.65	13.67	14.43	14.40	12.45	10.99
13.10	9.43	12.12	12.07	11.57	15.34	13.64	15.73	16.51	11.67
13.16	12.90	11.84	11.72	9.79	8.81	15.83	15.91	9.54	15.05
9.36	11.04	10.31	13.93	9.65	17.26	12.52	10.75	10.13	10.53
13.51	10.75	11.15	12.32	13.45	13.53	14.02	16.12	12.56	13.04
13.96	14.54	10.50	8.41	8.03	9.82	12.42	15.16	12.42	13.94
11.81	11.93	15.36	14.90	12.81	11.55	15.70	14.04	9.90	12.86
14.24	9.25	14.85	10.15	10.16	15.61	12.16	11.25	17.93	10.17
11.82	10.83	13.74	9.80	13.18	11.17	9.41	11.29	13.74	13.99
16.00	12.65	13.19	10.83	13.10	10.14	15.03	12.54	11.92	13.81
12.42	10.32	9.45	12.74	8.68	9.95	13.21	9.47	16.02	12.18
11.38	14.03	10.97	13.90	13.24	12.39	8.95	15.58	11.13	11.15
11.88	15.04	14.50	13.28	14.18	15.81	11.38	10.07	15.75	11.17
12.348	**11.732**	**11.748**	**11.798**	**11.558**	**13.189**	**13.266**	**12.657**	**12.909**	**12.439**

ANOVA 1: Every bivalve considered a true replicate

Source	Sum of Squares	DF	Mean Square	F-ratio	P
Distance	492.34	3	164.11	39.46	<0.0001
Error	1647.04	396	4.16		

ANOVA 2: Cages considered to be replicates; individual bivalves considered to be subsamples

Source	Sum of Squares	DF	Mean Square	F-ratio	P
Distance	492.34	3	164.11	62.40	<0.0001
Cages Within Distance*	42.13	16	2.63		
Clams Within Cages (= Subsampling Error)	1604.91	380	4.22		

* Cages within distance is the appropriate error term in the design with subsamplin

ANOVA 3: Cages considered to be replicates; mean values for growth within cages used in the analysis

Source	Sum of Squares	DF	Mean Square	F-ratio	P
Distance	24.62	3	8.21	62.40	<0.0001
Error	2.11	16	0.13		

The point of this discussion is to demonstrate that for the EEM studies, it may be advantageous to deploy replicate cages containing multiple bivalves at each station, and to consider only the average performance within each cage. This approach confers the additional advantage of providing redundancy in case one or more cages are lost, and may simplify the construction and deployment of cages. So long as the statistical hypothesis testing is confined to evaluating whether there are significant differences between areas (without uniquely attributing effects to effluent exposure), the replication and statistical analysis is valid.

As general guidance, it is suggested that cages of 20 animals per cage be deployed for the survival endpoint and at least five cages be deployed at each station to evaluate the growth endpoint. These cages should be deployed on individual moorings, not five cages on one mooring, since the mooring is the most appropriate unit of statistical replication. However, practitioners should be encouraged to explore the power and robustness of potential study designs, using synthetic data (or empirical data where available) as an integral part of the study design process to determine the minimum number of animals required.

11.2.7.4 Timing and Duration

Timing of the studies should be such that:

- it coincides with a high growth period in natural populations so that growth is maximized and differences in growth rate among treatments may be more measurable; and
- it does not coincide with a spawning period if the test bivalves are adults.

The duration of exposure should be 60 to 90 days (see discussion in Salazar and Salazar 2000). This should provide sufficient time for effects to survival and growth to be manifested.

11.2.7.5 Modifying or Confounding Factors and Supplementary Measures

Results of caged bivalve studies will depend, at least in part, on natural factors such as temperature, food supply, other physical/chemical properties of the test environments, species selected, condition of test organisms, exposure method, and handling of test organisms. Exposure and reference areas should be as similar as possible with respect to the factors listed below to minimize confounding differences. It may be useful to measure some of the factors to assist in interpreting results. These factors are presented in brief below and reviewed in more detail in Section 6 of Salazar and Salazar (2000).

Life cycle of test organisms is critical to the success of the study. The most important consideration is spawning cycle because of possible interference with growth that can confound the interpretation of results. Tests should be conducted with individuals that are not likely to spawn during the exposure period. If immature bivalves are used, this would not be an issue.

Temperature of the receiving environment can affect chemical uptake, survival and growth. Water temperature may be monitored during the exposure period to determine possible influences on test results and to evaluate similarity in temperature between areas. Continuous temperature monitoring devices are inexpensive (approximately $100 to $120 each) and readily available. One monitoring device could be deployed at each station by attaching to a cage.

Lack of acclimation to deployment water quality conditions may cause interference. It may be necessary to acclimate test organisms gradually to deployment conditions, particularly if collection site conditions are markedly different from those at deployment areas.

Food supply is important because it affects survival, growth, condition and chemical uptake. Food availability may be sampled during the exposure period to aid in interpretation of results. Measurements could include one or more of the following (depending on feeding mechanism of the test species): chlorophyll a, particulate or organic carbon, suspended solids.

Current speed may have an effect on growth rate of bivalves. It is important that investigators have a general understanding of current patterns and velocities for each area and station to aid in interpreting results. Current meters may be considered for stations with simple flow regimes. Currents of stations with complex flow regimes might be difficult to assess using meters. In such situations, a qualitative understanding of currents is sufficient.

Salinity is of particular importance in estuarine areas, where salinity can range from 0 parts per thousand near the mouth of the river to 33 parts per thousand within a short distance. Salinity can also vary to the same extent at some locations during the tidal cycle. A species with a wide salinity tolerance range should be chosen for such situations, although there is no guarantee that survival will not be affected (see Grout and Levings 2000). Salinity at each station may be measured using an in situ monitor. Local conditions should be observed to determine if salinity fluctuations are likely.

Bivalve cages can become fouled with plant and animal growth in highly productive embayments. **Excessive fouling** can reduce or eliminate water flow through the cage material, resulting in highly stressful conditions to the test organisms. If such conditions are anticipated, the deployed cages should be examined for fouling during the exposure period, and fouling removed by hand or with the use of a stiff brush.

Variability in chemical concentrations in the water column as a result of intermittent discharges may make it difficult to estimate exposure concentrations. Monitoring of discharge volumes and concentrations of tracer chemicals in the discharge during the exposure period will help to estimate exposure and interpret results. Chemical concentrations in receiving water could be sampled, for example, at beginning and end of the test to confirm presence/absence of effluent and other chemical influences in the cage deployment areas. Guidance on tracers is provided in Section 2.2.2 of the pulp and paper Technical Guidance Document (Environment Canada 1998) and Section 4.4.1.1.2 of the metal mining Guidance Document (Environment Canada 2001).

11.2.7.6 Cage Designs

Cages with individual compartments are suggested so that each individual bivalve can be tracked through the study. The mesh size should be maximized to allow maximum water flow but small enough to contain the test animals. Individuals are assigned to compartments such that survival, growth and condition can be tracked on each bivalve.

A variety of cage designs are described in Salazar and Salazar (2000). A flat (i.e., two-dimensional) cage design, as shown in Appendix I, is suggested, as it is a convenient unit to work with a number of reasons (M. Salazar, pers. comm.), including:

- flat cages have a single layer of bivalves that is easy to handle during set-up, deployment and retrieval;
- flat cages can be easily adjusted to accommodate a variable number of bivalves; and
- flat cages are quickly and easily stacked in a minimal space during set-up and transit to the deployment sites.

PVC tubing is a convenient material to use for constructing cages. PVC should be water-supply grade and from a high quality source. PVC tubing should be soaked before use for at least 24 hours in flowing fresh or seawater to remove water-soluble and volatile chemicals. Alternative materials are described in Salazar and Salazar (2000: Section 9).

Individual test organisms are placed in the mesh bags and separated by using a plastic cable tie or other suitable tie. Sufficient space should be allowed in each compartment to permit test animals to grow during the exposure period.

Final cage dimensions depend on the size of the test organism and the number of organisms per cage. Typical mesh sizes for species like mussels and clams are 10 to 15 cm diameter with 5 mm mesh size. Each mesh bag should be long enough to accommodate the desired number of bivalves per bag plus sufficient material for attachment to the PVC frame. The length of the frame should be 5 cm longer than each mesh sock with bivalves, and the width should be 5 cm greater on each end than the adjacent mesh sock to the frame.

For freshwater clam deployments, a wide variety of cage designs have been used by investigators. It is the responsibility of the study designer to ensure that cage designs are appropriate to the test species and receiving environment.

11.2.7.7 Mooring Systems

The cages can be suspended in the water column by attaching them to mooring lines that have an anchor or weight on one end (e.g., iron chain links) and a surface or subsurface buoy attached to the other end. The following factors should be considered for deployment of cages (from Salazar and Salazar 2000):

- appropriate vessel to accommodate cages with associated mooring lines, floats and anchors;
- tidal height (to ensure cages are at the desired depth during both low and high tides);
- slope and nature of the bottom material (to ensure the cages do not slide down a steep slope during the exposure period); and
- boating activity and recreational activity in the vicinity of the cages.

A number of mooring configurations are described in Salazar and Salazar (2000: Section 9) and in the pilot studies (Applied Biomonitoring 2000; Andrews and Parker 1999). A mooring system for a cube-shaped cage (from Andrews and Parker 1999) is illustrated in Figure 6. It may be wise to include a retrieval line on each cage to ease recovery at the end of the exposure period.

11.2.8 Variables to be Measured

The test variables to be measured in caged bivalve studies for EEM studies are survival growth and condition, as described in more detail below. An example of a reporting format for recording survival and growth raw data and endpoints is provided in Table 9. Reproductive condition is not recommended at this time by the EEM Science Committee for the reasons outlined below.

The EEM Science Committee does not feel that, at this time, there are well-developed, standardized methods of evaluating the reproductive condition of bivalves. The EEM Science Committee feels that this is an area that requires further research and development before this type of endpoint can be incorporated into an EEM program. An example of reproductive condition is gonad index used by Mucklow (1996) on mussels (*Mytilus edulis* and *M. trossulus*) and also on mussels by the Gulf of Maine Council on the Marine Environment (GMCME 1992). Gonad index is calculated as the ratio between gonad dry weight and tissue dry weight (GMCME 1992).

Guidance can be found in Salazar and Salazar (2000: Section 11) on how to analyze bivalve tissues. Pulp and paper mills may be required to conduct a tissue analysis of dioxins and furans if:

a) the concentration of dioxins and furans in the effluent has exceeded the prescribed limits in *the Pulp and Paper Mill Effluent Chlorinated Dioxins and Furans*, or
b) effluent has contained measurable concentrations of 2,3,4,8-TCDD or of 2,3,7,8-TCDF within the meaning of *the Pulp and Paper Mill Effluent Chlorinated Dioxins and Furans*, or
c) an effect on fish tissue was reported in the most recent EEM cycle.

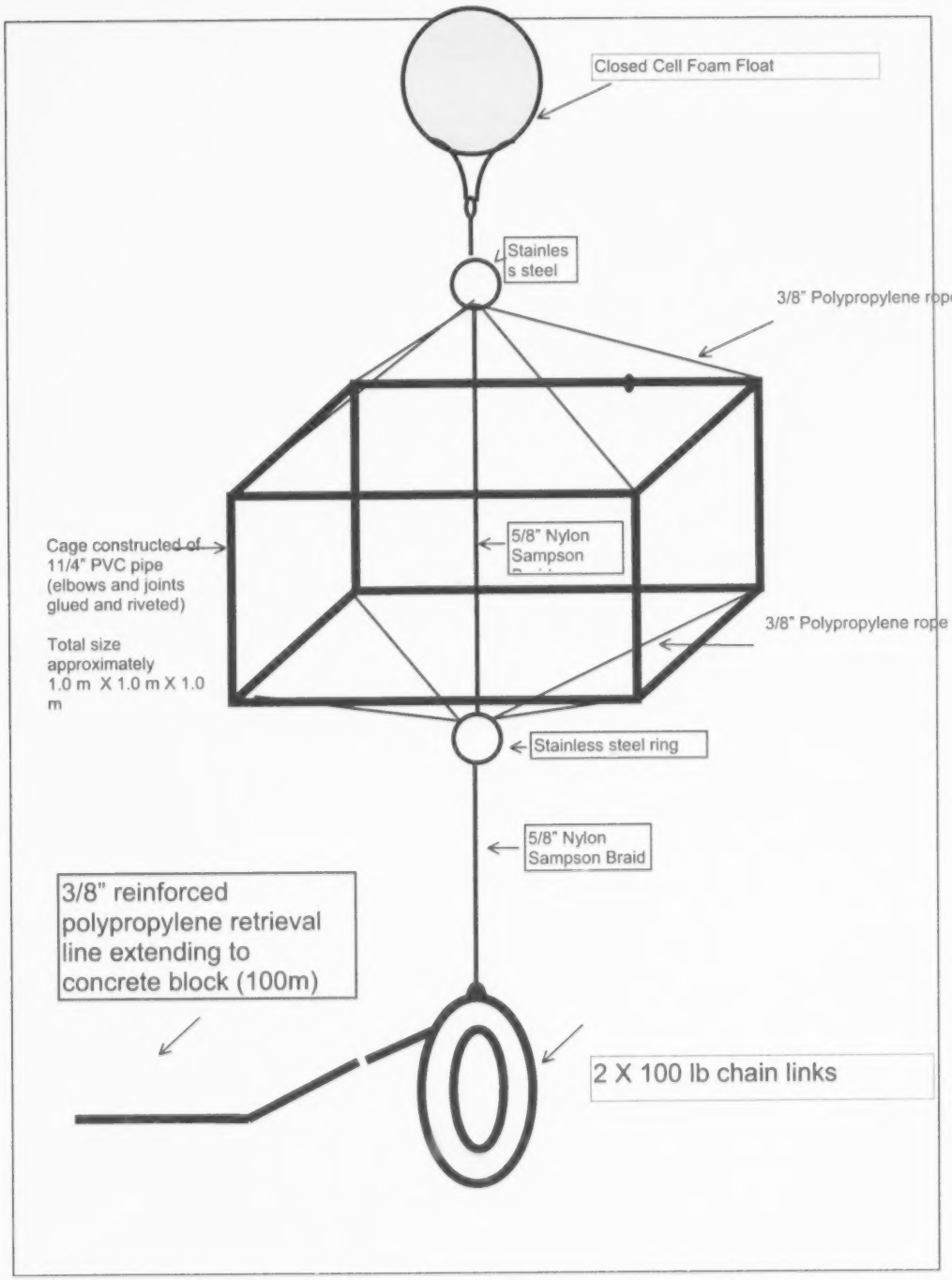


Figure 11-6: Example of bivalve mooring design

Table 11-9: An Example of A Field Data Sheet For Recording Survival And Growth Raw Data

Caged Bivalve Field Data Sheet

Study Area _____ Cage Identifier_____
Coordinates: Latitude _____N # of Mussels per cage
_____
Longitude _____W
Date/Time of Deployment _____ Date/Time of Retrieval _____
Days of Exposure _____ Depth of Water_____ Depth of Cage _____

Mussel Identifier	Length (0.01 mm)	Length (0.01 mm)	Whole Animal Wet Weight (0.01 g)	Whole Animal Wet Weight (0.01 g)	Soft Tissue Fresh Weight (0.01 g)
	Start	End	Start	End	End
1					
2					
3					
4					
5					
6					
7					
8					
9					
10					
11					
12					
13					
14					
15					
16					
17					
18					
19					
20					

Number of mussels alive at end of test _____ Percent Survival _____

Site Conditions:

Date	Time	Temperature (°C)	Dissolved Oxygen (mg/L)	Salinity (ppt)	Current Velocity (cm/sec)	Current Direction (degrees)

Field Staff: _____

Notes:
_____

11.2.8.1 Survival

Survival is not a particularly sensitive indicator of effects in caged bivalves, but is an important parameter to monitor. Survival can be easily determined and quantified, although it is possible to have some individuals missing at the test end due to shell decomposition. Bivalves are dead if they are gaping and do not close their shells when touched or tapped. Survival is expressed as percent of individual animals alive per cage at the end of the exposure period.

11.2.8.2 Growth

Growth is a measure of energy use and is a sensitive indicator of effects that is easy to measure. There are several types of growth measurements that should be made; measuring only one could provide misleading results. Growth measurements are outlined in Table 10.

Table 11-10: Growth Measurements

Growth Measurement	Measured at:		How Measured	Accuracy [2]
	Test Initiation	Test Termination [1]		
Whole animal wet weight (WAWW)	✓	✓	Analytical balance	0.01 g
Shell length	✓	✓	Vernier calipers; longest axis	0.01 mm
Soft tissue fresh weight	✓ (on sacrificed subsample)	✓	Analytical balance	0.01 g

[1] Test termination measurements should only be made on surviving individuals.

[2] Accuracy is based on larger bivalve species such as marine mussels and freshwater unionids. Accuracy may need to be adjusted downward for smaller-sized test species, such as sphaeriid clams.

Whole animal wet weight (WAWW) and shell length are non-destructive measurements and should be made on each individual at the beginning and end of the test. Soft tissue fresh weight should also be measured at the beginning and end of the test. However, as this is a destructive test, measurements at the beginning of the test should be made on a subsample of bivalves from the same population and size range as the test organisms. Soft tissue weight should be measured as fresh weight (e.g., tissue weight should be measured as soon as possible after test termination while the soft tissue is still fresh and has not been frozen or left to dry out). Weights (WAWW and soft tissue fresh weight) should be determined with an analytical balance with a measurement accuracy of 0.01g. Shell length should be determined with vernier calipers along the longest axis with a

measurement accuracy of 0.01 mm. Please note that accuracy of these measurements may need to be adjusted downwards for smaller-sized test species.

Growth can be expressed in a number of ways:

- absolute growth: = absolute change in value from test initiation to test termination;
- growth rate: = absolute change in value per unit time, typically using one week as the time increment; or
- relative growth: = (final weight – initial weight) / initial weight; relative growth may be used when there is a significant difference among cages in initial weights; relative growth is expressed as a proportion and therefore an arcsin square root transformation of relative growth values may be appropriate prior to applying statistics. Green (1979: Page 52) provides useful advice on this transformation, noting that it is not usually required for proportion data in the range of 0.3 to 0.7, and that while it may not always help, it probably does no harm either.

Use the most appropriate expression of growth to suit the study design and site-specific characteristics.

11.2.8.3 Condition

Condition is a measure of how an animal stores its energy. There are more than one option that may be considered for calculating condition, as described below. Note that some of these methods require measurement of additional variables to those outlined in Section 11.2.8.2. The most appropriate method to calculate condition is left to the discretion of the investigators.

Weight (WAWW, dry shell or soft tissue weight) related to shell length

This is analagous to the Fulton Condition Index (Ricker 1975; Anderson and Neumann 1996) used in fisheries biology. This relationship may be characterized according to a conventional formula for a straight line (e.g., in Mackie and Flippance 1983), with slope (C) and intercept (b):

$$\log W = b + (C \times \log L)$$

High values of C imply that a bivalve has relatively high tissue weight at a given length, whereas low values may indicate that animals are not obtaining sufficient food, or are experiencing chronic stress that prevents them from thriving. This method of characterizing condition is suitable for assessing condition in wild bivalves where shell length is expected to be quite variable. However, since animals used for caged bivalve studies are screened for uniform length at test initiation, this method will not be reliable. This method may also not be suitable for bivalves because shell length and tissue weight are influenced by different factors in the environment (M. Salazar, pers. comm.).

Soft tissue weight related to shell weight

This method of characterizing condition uses soft tissue weight and shell weight. An ANOVA can be conducted or, more simply, soft tissue weight can be divided by shell weight. Grout and Levings (2000) measured condition of blue mussels as the ratio of tissue weight to shell weight. They found that condition distinguished caged mussels in a high survival zone (condition index 1.10 to 1.42) from caged mussels in a low survival zone (condition index 0.82 to 0.96).

Soft tissue weight related to shell volume

This method was used by Mucklow (1996; based on Seed 1968) to calculate condition by dividing soft tissue dry weight by shell volume, measured as length x width x height. In a study on wild blue mussel populations, Mucklow (1996) concluded that seasonal patterns in condition index were variable and influenced by a number of natural factors, including food availability and physiological energy requirements.

11.2.9 Methods for Test Initiation, Cage Deployment and Retrieval, Test Termination

These methods are described in detail in Salazar and Salazar (2000: Section 11) and summarized below.

11.2.9.1 Test Initiation

The first step is to sort all bivalves into the desired size range(s). By selecting bivalves within a narrow size/age range, the investigator can be relatively confident that the individuals will have similar growth potential. Determining the age of wild-collected bivalves may be difficult if not impossible for some species; length is therefore used to obtain individuals with similar growth rates. Commercial growers can often provide bivalves of a known age.

The size range selected for test organisms depends on the test species, test organisms supply, and target age. Test organisms should be selected within a narrow size range. Some studies have selected bivalves with a narrow size range and randomly allocated them among cages (e.g., Andrews and Parker 1999). Other studies have used multiple size classes and allocated bivalves to cages such that there is no significant difference among cages in terms of initial size and weight (see Salazar and Salazar 2000: Section 11).

Test organisms should be kept cool and moist during the sorting stage to prevent stress, damage or death (see Section 11.2.4.7). Shell lengths and WAWW are recorded for each

animal and the animals assigned to individual holding containers (e.g., ice cube tray) that are labelled according to specific location within the cage. This allows measurements on each individual test organism to be repeated at test termination. Animals in a sacrificed subsample (statistically appropriate number) of test organisms will be measured for shell length, WAWW, tissue weight and shell weight.

For quality control purposes, approximately 10 percent of the measurements should be repeated and recorded by a different investigator. It is best to use an electronic spreadsheet for recording all measurements to reduce transcription errors. It is possible to electronically link measurement tools (e.g., vernier calipers and analytical balance) with a spreadsheet to provide an additional quality control measure.

All test organisms needed to fill the cages, plus a subsample for destructive sampling (tissue weight), should be sorted and measured prior to distribution to socks and cages. Test organisms are then distributed to individual mussel socks before they are attached to cages. An example of a labeling and distribution scheme is provided in Salazar and Salazar (2000: Section 11).

To ensure statistical similarity in WAWW and shell length among stations, an ANOVA can be conducted. This can be programmed into a spreadsheet such that the data are analyzed by cage and by station (if more than one cage is to be deployed at each station). If statistically significant differences are found, the test animals should be redistributed to eliminate this difference. This may require opening mesh socks and re-closing with a plastic tie.

11.2.9.2 Cage Deployment

Ensure that an appropriate vessel is chartered to ensure safe transport and deployment of all cages, moorings and floats. Moorings and floats may be attached to the cages on the vessel on route to site. When on site, cage location should be identified using Global Positioning System (GPS), or some other reliable method (e.g., nearby on-shore reference points). During the exposure period, the cages may be inspected (e.g., using SCUBA divers) for presence, damage and fouling.

11.2.9.3 Cage Retrieval and Test Termination

Cages should be retrieved using various location identifiers (e.g., GPS, depth sounders) and the use of a grappling hook if a retrieval rope was used for each cage.

Details on test termination are provided in Salazar and Salazar (2000: Section 11) and outlined here. At the processing site, bivalves from all bags for each cage should be processed together. It is essential that the order and orientation of each bivalve is

maintained during all of the end-of-test measurements. Individuals can be removed from mesh socks into labelled individual containers (e.g., ice cube trays) to facilitate measurement.

At 5 to 10 minutes prior to making length and weight measurements, set the tray(s) into a tub containing clean water. Individuals that float indicate that air is trapped in between the valves of the shell. When all animals in the tray(s) have opened their valves slightly and are not longer floating (5 to 10 minutes), length and weight measurements may begin. Start with WAWW and shell length measurements, then carefully shuck each individual and blot the tissue before making tissue fresh weight measurements. Note that if tissues are to be analyzed for chemical parameters, then care must be taken not to introduce any contaminants from the blotting material.

11.2.10 Data Analyses

Data analyses and interpretation of the results should be appropriate for the study design. Caged mussel studies for EEM are designed to determine if there are significant effects on biota in the vicinity of effluent outfalls. This can be accomplished by using either a control-impact (= reference-exposure) or a gradient (= regression) type of design, as discussed in Section11.2.5 Statistical procedures appropriate for each effect endpoint (i.e., survival, growth, and condition) are summarized in Table 11.

Table 11-11: Summary of Statistical Procedures to Analyze Effect Endpoints for Caged Bivalve Studies

Type of Response	Effect Endpoint	Statistical Procedures
Survival	Survival	ANOVA
Energy Use	Growth – based on shell length	ANOVA (regression analysis for gradient designs)
	Growth – based on WAWW	ANOVA (regression analysis for gradient designs)
	Growth – based on soft tissue fresh weight	ANOVA (regression analysis for gradient designs)
Energy Storage	Condition	ANOVA or ANCOVA
Bioaccumulation*	Tissue Chemistry*	Multivariate statistics

* May not be required for EEM (determined on a site specific basis), but may be used to assist in data interpretation.

The first step is to generate summary statistics for each parameter (i.e., WAWW, shell length, soft tissue fresh weight) on a cage and station basis.

The second step is to determine whether there are significant differences among replicate cages for each of the parameters measured before deployment (if they were not measured post-distribution and before deployment). This involves assessing the data for normality and homogeneity of variances.

The final step is to use the appropriate statistical test (see Table 11) for the study design. In general, ANOVA and multiple comparison tests are used for hypothesis testing and comparison among stations. For C-I and MC-I designs using ANOVA and Analysis of Co-Variance (ANCOVA) procedures, detailed guidance is provided in Sections 5.1.5.3 to 5.1.5.5 and in 5.1.6.1 of the pulp and paper Technical Guidance Document (Environment Canada 1998) and Section 5.16 of the metal mining Guidance Document (Environment Canada 2001). If statistical differences are found, a multiple range test, or its non-parametric counterpart, can be used to determine which stations are different from the others. Linear and multiple regression analyses (using variables regressed on distance) may generally be used to establish relationships among variables along exposure gradients. For ANCOVA analyses of condition, the covariates will be dictated by the condition formula that is used.

11.2.11 Tissue Concentrations

11.2.11.1 Dioxins and Furans

If a mill is required to measure dioxins and furans in fish tissue, then measuring these chemicals in caged bivalves may be considered. However, there are some precautions that should be taken with this approach, including the following:

- duration of exposure should be chosen such that dioxins and furans are likely to be accumulated to detectable levels in bivalve tissue;
- consider whether bivalves would be harvested (commercially or recreationally) for human consumption in the area; and
- ensure that there are sufficient number of bivalves included in the design to obtain enough tissue for analysis.

There may be mills for which caged bivalves would not be considered suitable for measuring uptake of dioxins and furans. Any proposal to use caged bivalves for dioxin and furan measurements would require approval by the Regional Authorization Officer.

11.2.11.2 Other Chemicals

If there is a requirement for a particular mill to measure tissue concentrations to assess bioavailability of particular chemicals, then refer to guidance in Salazar and Salazar (2000: Section 11) on sampling design considerations and handling.

11.2.12 Reporting

Quality assurance and quality control (QA/QC) considerations for caged bivalve studies should follow those outlined for the Fish Survey in Section 5.1.8 of the pulp and paper Technical Guidance Document (Environment Canada 1998) and Section 4.14 of the metal mining Guidance Document (Environment Canada 2001). QA/QC measures apply to the following components of caged bivalve studies:

- study design;
- field sampling;
- sample processing/laboratory analysis;
- data analysis; and
- reporting.

11.2.13 References

Anderson, R.O. and R.M. Neumann. 1996. Length, weight, and associated structural indices. Pp. 447-482. In: B.R. Murphy and D.W. Willis (eds.). Fisheries Techniques. Second Edition. American Fisheries Society, Bethesda, Maryland.

Andrews, S. and R. Parker. 1999. Draft Report (II). An Environmental Quality Evalutaion of Pictou Harbour, Nova Scotia, Using Caged Bivalves; *Mytilus edulis*. Prepared by Environmental Protection Branch of Environment Canada (Dartmouth, NS) for The Pictou Harbour Environmental Protection Project, New Glasgow, NS.

Applied Biomonitoring. 2000. Caged Mussel Pilot Study Port Alice Mill, Vancouver Island EEM Project. Prepared for Pacific and Yukon Region of Environment Canada, North Vancouver, BC.

BC Research Inc. 1998. Methyl Mercury Impact Assessment for Nechako Reservoir Submerged Timber Salvage Operations: Caged Mussel Study. Prepared for Ministry of Environment, Lands and Parks, Environmental Protection – Skeena Region, Environmental Impact Section, Smithers, BC.

Beckvar, N., S. Salazar, M. Salazar and K. Finkelstein. 2000. An in situ assessment of mercury contamination in the Sudbury River, Massachusetts, using transplanted freshwater mussels (*Elliptio complanata*). Can. J. Fish. Aquat. Sci. 57: 1103-1112.

Clarke, A.H. 1973. The freshwater molluscs of the Canadian interior basin. Malacologia 13, (1-2): 1-509.

Clarke, A.H. 1981. The freshwater molluscs of Canada. National Museum of Natural Sciences, National Museums of Canada, Ottawa, ON. 446 pp.

Couillard, Y., P.G.C. Campbell, A. Tessier, J. Pellerin-Massicotte and J.C. Auclair. 1995a. Field transplantation of a freshwater bivalve, *Pyganodon grandis*, across a metal contamination gradient. I. Temporal changes in metallothionein and metal (Cd, Cu, and Zn) concentrations in soft tissues. Can. J. Fish. Aquat. Sci., 52: 690-702.

Couillard, Y., P.G.C. Campbell, A. Tessier, J. Pellerin-Massicotte and J.C. Auclair. 1995b. Field transplantation of a freshwater bivalve, *Pyganodon grandis*, across a metal contamination gradient. II. Metallothionein response to Cd and Zn exposure, evidence for cytotoxicity, and links to effects at higher levels of biological organization. Can. J. Fish. Aquat. Sci., 52: 703-715.

Courtenay, S.C., W.R. Parker and G.P. Rawn. 1998. Proceedings of a workshop to assess alternatives to the fish survey component of the environmental effects monitoring program for Canadian pulp and paper mills. Can. Tech. Rep. Fish. Aquat. Sci., 2233: viii + 108 pp.

Day, K.E., J.L. Metcalfe and S.P. Batchelor. 1990. Changes in intracellular free amino acids in tissues of the caged mussel, *Elliptio complanata*, exposed to contaminated environments. Arch. Environ. Contam. Toxicol., 19: 816-827.

Environment Canada. 2001. Metal Mining Guidance Document for Aquatic Environmental Effects Monitoring. Environment Canada. Hull, Québec (draft).

Environment Canada. 1998. Pulp and Paper Technical Guidance for Aquatic Environmental Effects Monitoring EEM/1998/1.

Freeman, K.R., K.L. Perry, T.G. DiBacco and D.J. Scarratt. 1994. Observations on two Mytilid species from a Nova Scotian mussel farm. Can. Tech. Rep. Fish. Aquat. Sci. No. 1969.

GMCME (Gulf of Maine Council on the Marine Environment). 1992. Evaluation of Gulfwatch 1991 Pilot Project of the Gulf of Maine Marine Environmental Monitoring Plan. Gulf of Maine Council on the Marine Environment, State Planning Office, August, ME.

Goldberg, E.D., V.T. Bowen, J.W. Farrington, G. Harvey, J.H. Martin, P.L. Parker, R.W. Risebrough, W. Robertson, E. Schneider and E. Gamble. 1978. The Mussel Watch. Environ. Conserv., 5(2): 101-125.

Green, R.H. 1979. Sampling Design and Statistical Methods for Environmental Biologists. John Wiley & Sons, Toronto, ON. 257 pp.

Grout, J.A. and C.D. Levings. 2000. Effects of acid mine drainage from an abandoned copper mine, Britannia Mines, Howe Sound, British Columbia, Canada, on transplanted blue mussels (*Mytilus edulis*). Marine Environmental Research, 51(3): 265-288.

Hinch, S.G. and R.H. Green. 1989. The effects of source and destination on growth and metal uptake in freshwater clams reciprocally transplanted among south central Ontario lakes. Can. J. Zool., 67: 855-863.

Hornbach, D.J., T.E. Wissing and A.J. Burky. 1982. Life-history characteristics of a stream population of the freshwater clam *Sphaerium striatinum* Lamarck (Bivalvia: Pisidiidae). Can. J. Zool., 60: 249-260.

Hurlbert, S.H. 1984. Pseudoreplcation and the design of ecological field experiments. Ecol. Monog., 54: 187-211.

Jernelov, A. 1996. The International Mussel Watch: A global assessment of environmental levels of chemical contaminants. Sci. Tot. Environ., 188 (Suppl. 1): 37-44.

Mackie, G.L. 1978a. Effects of pollutants on natality of *Musculium securis* (Bivalvia: Pisidiidae). The Nautilus, 92 (1): 25-33.

Mackie, G.L. 1978b. Are sphaeriid clams ovoviviparous or viviparous? The Nautilus, 92: 145-147.

Mackie, G.L., S.U. Qadri and A.H. Clarke. 1974. Development of brood sacs in *Musculium securis* Bivalvia: Sphaeriidae. The Nautilus, 88: 109-111.

Mackie, G.L. and L.A. Flippance. 1983. Growth dynamics in *Sphaerium rhomboideum* (Bivalvia: Pisidiidae). Can. J. Zool., 61: 868-873.

Malley, D.F., A.R. Stewart and B.D. Hall. 1996. Uptake of methyl mercury by the floater mussel, *Pyganodon grandis* (Bivalvia, Unionidae) caged in a flooded wetland. Environ. Toxicol. Chem., 6: 928-936.

McMahon, R.F. 1991. Mollusca: Bivalvia. In J.H. T.G. Horp and A.P. Covich (eds). Ecology and Classification of North American Freshwater Invertebrates. Academic Press Inc., San Diego, California. Pp. 315-399.

Metcalfe-Smith, J.L., R.H. Green and L.C. Grapentine. 1996. Influence of biological factors on concentrations of metals in the tissues of freshwater mussels (*Elliptio complanata* and *Lampsilis radiata radiata*) from the St. Lawrence River. Can. J. Fish. Aquat. Sci., 53: 205-19.

Mucklow, L.C. 1996. Effects of season and species on physiological condition and contaminant burdens in mussels (*Mytilus edulis L.* and *Mytilus trossulus G.*):

implications for mussel watch programs. Thesis submitted in partial fulfillment of requirements for Master of Environmental Studies, School for Resource and Environmental Studies, Dalhousie University, Halifax, NS.

Newell, R.I.E. 1989. Species Profiles: Life Histories and Environmental Requirements of Coastal Fishes and Invertebrates (North and Mid-Atlantic) – Blue Mussel. U.S. Fish Wildl. Serv. Biol. Rep. 82 (11.102). US. Army Corps of Engineers, TR EL-82-4.

Nichols, S.J. 2000. Nonindigenous Species Research and Outreach. Life History and Ecological Requirements of the Zebra Mussel – North American Experience Through 1992. http://www.nsgo.seagrant.org/research/nonindigenous /zmlifehistory.html.

O'Connor. 1992. Recent trends in coastal environmental quality: Results from the first five years of the NOAA Mussel Watch Project. Washington, DC. U.S. Department of Commerce, National Oceanic and Atmospheric Administration (NOAA), National Ocean Service.

Paine, M.D. 2000. EEM Design. Short-Course presented at the 27th Annual Aquatic Toxicity Workshop, October 1-4th, St. John's, NF.

Ricker, W.E. 1975. Computation and interpretation of biological statistics of fish populations. Bulletin 191. Department of the Environment and Fisheries and Marine Service. Ottawa, ON.

Salazar, M. 2001. Applied Biomonitoring, Seattle, WA. Personal Communication via conference call, March 2001.

Salazar, M.H. and S.M. Salazar. 1995. In situ bioassays using transplanted mussels: I. Estimating chemical exposure and bioeffects with bioaccumulation and growth. Pp: 216-241. In: J.S. Hughes, G.R. Biddinger and E. Mones (eds). Environmental Toxicology and Risk Assessment – Third Volume. American Society for Testing and Materials (ASTM), Philadephia, PA.

Salazar, M. and S. Salazar. 2000. Standard Guide for Conducting In-Situ Field Bioassays with Caged Marine, Estuarine and Freshwater Bivalves. (Draft dated November 17, 2000). American Society for Testing and Materials (ASTM), Philadelphia, PA. Accepted for publication in 2001.

Sandusky, M.J. and R.E. Sparks. 1979. Investigations of declines in fingernail clam (Musculium transversum) populations in the Illinois River and Pool 19 of the Mississippi River. Bull. Amer. Malacol. Union Inc., 1979: 11-16.

Seed, R. 1968. Factors influencing shell shape in *Mytilus edulis* L. J. Mar. Biol. Ass. U.K., 48: 561-584.

Skidmore, D. and K.K. Chew. 1985. Mussel aquaculture in Puget Sound. Washington Sea Grant Publication, Seattle, Washington.

Stephenson, M. and G.L. Mackie. 1981. Effects of 2, 4-D treatment on growth and survival of fingernail clams (Bivalvia: Pisidiidae) in artificial pond ecosystems. Bull. Amer. Malacological Union, 1981: 19-22.

Stewart, P.L. 1994. Environmental requirements of the blue mussel (*Mytilus edulis*) in Eastern Canada and its response to human impacts. Can. Tech. Rep. Fisheries Aquat. Sci., 2004. 41 pp.

Stewart, R. and D.F. Malley. 1997. Technical evaluation of molluscs as a biomonitoring tool for the Canadian mining industry. Aquatic Effects Technology Evaluation (AETE) Program Project 2.3.1. Sponsored by Canada Center for Mineral and Energy Technology (CANMET) and the Mining Association of Canada (MAC).

Virginia Tech. 2000. Fish and Wildlife Information Exchange, Conservation Management Institute, Species Profile. Blacksburg, VA. (**Error! Bookmark not defined.**, Virginia Tech.

Waldock, M.J., J.E. Thain and M.E. Waite. 1996. An assessment of the value of shell thickening in *Crassostrea gigas* as an indicator of exposure to tributyltin. Pp. 219-237. In: M.A. Champ and P. F. Seligman (eds). Tributyltin: Environmental Fate and Effects. Chapman and Hall, London.

Williams, J.D., M.L. Warren Jr., K.S. Cummings, J.L. Harris and R.J. Neves. 1993. Conservation status of freshwater mussels of the United States and Canada. Fisheries, 18 (9): 6-22.